

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-232838

(43)Date of publication of application : 16.08.2002

(51)Int.Cl. H04N 5/91

G03B 19/02

H04N 5/225

H04N 5/232

H04N 5/262

H04N 13/02

// H04N101:00

(21)Application number : 2001-110393 (71)Applicant : MONOLITH CO LTD

(22)Date of filing : 09.04.2001 (72)Inventor : AKIYOSHI GIYOUZO
AKIYOSHI NOBUO
SHINAGAWA YOSHIHISA

(30)Priority

Priority number : 2000360638

Priority date : 28.11.2000

Priority country : JP

(54) DIGITAL CAMERA

(57)Abstract:

PROBLEM TO BE SOLVED: To solve the problem of the compressibility of moving pictures caused by the presence of the limit of a memory capacity in a digital camera.

SOLUTION: An imaging part 52 inputs the first image and the second image. A matching processor 14 performs pixel matching between the pictures to obtain the corresponding point of the grid points of a mesh arranged on the first picture on the second picture. The result is recorded as a corresponding point file F. An intermediate picture generation part 18 generates the intermediate picture between the first picture and the second picture based on the file F. This approximates the moving pictures. Since the corresponding point file is generated by the grid points, the data capacity is very small.

* NOTICES *

JP0 and INPIT are not responsible for any
damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1]A digital camera comprising:

An image pick-up part which photos a picture.

A camera controller which makes an image pick-up part photo the 1st and 2nd pictures at the predetermined intervals, and a matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs the result as a corresponding-points file.

[Claim 2]A digital camera comprising:

An image pick-up part which photos a picture.

A camera controller which determines a picture of two sheets as the 1st and 2nd pictures among pictures photoed by image pick-up part, and a matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs

the result as a corresponding-points file.

[Claim 3]A digital camera comprising:

An image pick-up part which realizes stereo **.

A camera controller which makes an image pick-up part photo the 1st and 2nd pictures that constitute a stereo image, and a matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs the result as a corresponding-points file.

[Claim 4]The digital camera according to any one of claims 1 to 3 which contains further an intermediate image generating part which generates an intermediate image of the 1st picture and the 2nd picture based on said corresponding-points file.

[Claim 5]The digital camera according to claim 4 which contains further an indicator which displays the 1st picture, the 2nd picture, and an intermediate image.

[Claim 6]A digital camera given in either of claims 4 and 5 which contain further a corresponding-points file attaching part which associates and records the 1st picture, the 2nd picture, and a corresponding-points file.

[Claim 7]Said matching processor detects a point on the 2nd picture corresponding to a lattice point of a mesh provided on the 1st picture by image matching, The digital camera according to any one of claims 1 to 6 which defines a terminal point polygon corresponding to a starting point polygon which constitutes a mesh on the 1st picture on the 2nd picture based on the result.

[Claim 8]The digital camera according to claim 7 in which said matching processor performs matching calculation of a pixel unit based on correspondence with a singular point which detected by performing two-dimensional search about the 1st picture, and a singular point which detected by performing two-dimensional search about the 2nd picture.

[Claim 9]After said matching processor forms the 1st picture and the 2nd picture into multiplex resolution by extracting said singular point, respectively, The digital camera according to claim 8 which acquires a correspondence relation of a pixel unit in a level whose resolution is final the finest while performing matching calculation of a picture unit between the same resolution levels and inheriting the result to matching calculation in a different resolution level.

[Claim 10]The digital camera according to claim 1 which has further a mode setting part for setting up the mode of simple animation photography by said image pick-up part.

[Claim 11]Including an image pick-up part which acquires the 1st picture and the 2nd picture, and a matching processor which calculates matching of the 1st picture and the 2nd picture, this matching processor, A digital camera generating a corresponding-points file by defining a terminal point polygon corresponding to a

starting point polygon which constitutes a mesh on the 1st picture on the 2nd picture. [Claim 12]An image pick-up part which photos a picture, and a camera controller which makes an image pick-up part photo the 1st and 2nd pictures at the predetermined intervals, Including a matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs that result as a corresponding-points file, this matching processor, A digital camera detecting correspondence of the 1st picture and the 2nd picture located in the final finest hierarchy by hierarchizing each of the 1st picture and the 2nd picture by extracting a singular point, and detecting a correspondence relation of singular points from the hierarchy's coarser one one by one.

[Claim 13]An image pick-up part which photos a picture, and a camera controller which determines a picture of two sheets as the 1st and 2nd pictures among pictures photoed by image pick-up part, Including a matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs that result as a corresponding-points file, this matching processor, A digital camera detecting correspondence of the 1st picture and the 2nd picture located in the final finest hierarchy by hierarchizing each of the 1st picture and the 2nd picture by extracting a singular point, and detecting a correspondence relation of singular points from the hierarchy's coarser one one by one.

[Claim 14]An image pick-up part which realizes stereo **, and a camera controller which makes an image pick-up part photo the 1st and 2nd pictures that constitute a stereo image, Including a matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs that result as a corresponding-points file, this matching processor, A digital camera detecting correspondence of the 1st picture and the 2nd picture located in the final finest hierarchy by hierarchizing each of the 1st picture and the 2nd picture by extracting a singular point, and detecting a correspondence relation of singular points from the hierarchy's coarser one one by one.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention]Especially this invention relates to the digital camera which processes predetermined in the photoed picture about a digital camera.

[0002]

[Description of the Prior Art]BS digital broadcasting is started, and many users come to enjoy the service on the Internet from PC or a gestalt telephone, and the fence of

broadcast and communication is collapsing quickly. Video equipment and a camera digitize and an individual youth's information machines and equipment came to have broadcast, communication, and very high compatibility. The "multimedia" said for a concept not to be in sight easily temporarily became one driving force of human-beings culture by technical innovation and maintenance of the infrastructure today when the 21st century came around.

[0003]The digital camera which regarded the efficient preservation and printing by digitization of a picture at the beginning, and appeared also carries various image processings technique nowadays, and the function to also bear a professional youth with strengthening of a personal color is realized. There is no room for the digital instrument led by an individual to put and doubt accelerating progress of IT culture from now on.

[0004]

[Problem(s) to be Solved by the Invention]The art which generates the edge enhancement by a highpass filter and the picture suitable for the user's taste, such as a color tone conversion process, is used for some digital cameras. Art, such as compression by MPEG (Motion Picture Expert Group), is adopted that an animation should be photoed if possible for a long time. It is in such a situation and the purpose of this invention is in offer of the digital camera which realizes an animation by comparatively light data by adoption of new image processing.

[0005]

[Means for Solving the Problem]A mode with this invention is related with a digital camera. This camera uses image matching about time. That is, this camera calculates matching between an image pick-up part which photos a picture, a camera controller which makes an image pick-up part photo the 1st and 2nd pictures at the predetermined intervals, and the 1st picture and the 2nd picture, and contains a matching processor which outputs that result as a corresponding-points file. A user sets up "a predetermined interval" and it may be fixed beforehand.

[0006]A camera controller photos a picture of two sheets in order an interval predetermined [this], when a user points to photography of a picture of one sheet, for example. Since a matching processor makes a corresponding-points file based on matching of a picture of two sheets, this file is behind used for it and it can generate an intermediate image of a picture of two sheets. As a result, an animation is reproducible in simple with small data volume. If a photographing interval of a picture of two sheets is opened to some extent, a morphing picture rather as used in the field of an image effect will be acquired rather than reproduction of an animation. It is also dramatically interesting as a function of a camera. For example, if a picture of two sheets is used as two persons' another face, a morphing picture of those faces is generable.

[0007]Another mode of this invention is also related with a digital camera. This

camera is provided with the following.

An image pick-up part which photos a picture.

A camera controller which determines a picture of two sheets as the 1st and 2nd pictures among pictures photoed by image pick-up part.

A matching processor which calculates matching between the 1st picture and the 2nd picture, and outputs the result as a corresponding-points file.

A camera controller may determine a predetermined picture of two sheets as the 1st and 2nd pictures in accordance with a predetermined rule among pictures which an image pick-up part photoed according to a user's directions. According to this mode, an above-mentioned morphing picture or a compression animation can be obtained with higher flexibility. This mode uses image matching about time, space, or both.

[0008]Another mode of this invention is also related with a digital camera. This camera uses image matching about space. Namely, an image pick-up part in which this camera realizes stereo ** and a camera controller which makes the 1st and 2nd pictures from which a stereo image is constituted by an image pick-up part photo. Matching is calculated between the 1st picture and the 2nd picture, and a matching processor which outputs the result as a corresponding-points file is included. Therefore, a special effect image and a viewpoint change picture are generable based on this corresponding-points file. It is for depth information of each point of the picture to become clear based on corresponding points of a stereo image.

[0009]This camera may also contain further an intermediate image generating part which generates an intermediate image of the 1st picture and the 2nd picture based on a corresponding-points file. An intermediate image is an interpolation picture concerning the both by time or space, and a case. It may have further an indicator which displays the 1st picture, the 2nd picture, and an intermediate image as an animation, a middle visual point picture, etc. A corresponding-points file attaching part which associates and records the 1st picture, the 2nd picture, and a corresponding-points file, for example, an IC card and other memory cards, and those control circuits may also be included further.

[0010]A matching processor may detect a point on the 2nd picture corresponding to a lattice point of a mesh provided on the 1st picture by image matching, and may define a terminal point polygon corresponding to a starting point polygon which constitutes a mesh on the 1st picture in the 2nd picture top based on the result. A matching processor may perform matching calculation of a pixel unit between the 1st picture and the 2nd picture. This calculation may follow all the pixels and may follow a lattice point or a lattice point, and a pixel relevant to it.

[0011]A matching processor may perform matching calculation of a pixel unit based on correspondence with a singular point which detected by performing two-dimensional search about the 1st picture, and a singular point which detected by performing two-dimensional search about the 2nd picture. In this case, after forming

the 1st picture and the 2nd picture into multiplex resolution by extracting a singular point, respectively further, Matching calculation of a picture unit is performed between the same resolution levels, and a correspondence relation of a pixel unit in a level whose resolution is final the finest may be acquired, inheriting the result to matching calculation in a different resolution level.

[0012]These people are application of art (henceforth "premise art") previously proposed by the patent No. 2927350, and are suitable for a matching method using the above-mentioned singular point to processing of a matching processor. However, premise art is not touching at all the feature of this invention about a polygon decided by a lattice point or it. Since a kind of simplification art of a polygon was introduced by this invention, size of said corresponding-points file can be made very small.

[0013]That is, if correspondence of those pixel units is described as it is when the 1st and 2nd picture has a $n \times m$ pixel, respectively, combination of $^2(n \times m)$ will arise and a corresponding-points file will become huge data volume. However, although it is a correspondence relation of lattice points, or the substantially same thing about this, data volume becomes less substantially by changing into describing a correspondence relation of the polygons decided by a lattice point. If only the 1st and 2nd picture and a corresponding-points file are held, an animation can be reproduced and a big effect will be demonstrated in transmission of an animation, storage, etc. It is the suitable art for a digital camera which has restriction in storage capacity of a picture.

[0014]In premise art, this invention does not have an end as it is indispensable. Replace each above composition and a process arbitrarily, or between a method and a device, in part or all replace expression, it is added, or what changed expression into a computer program, a recording medium, etc. is effective as this invention.

[0015]

[Embodiment of the Invention]First, the image matching processing using the multiplex resolution singular point filtering technique and it which are used by an embodiment is explained in full detail as "premise art." Such art is the art in which these people have already acquired the patent No. 2927350, and is the best for combination with this invention. Since a mesh is provided on a picture and many pixels are represented with this invention by the lattice point, it is because the application effect over the matching art of a pixel unit like premise art from the first is high. However, image matching art employable by an embodiment is not restricted to this. The digital camera using premise art is concretely explained after drawing 18.

[0016]The [background of premise art] The component engineering of premise art explains in full detail by [1] first, and [2] explains procedure concretely. Furthermore, [3] reports the result of an experiment.

[0017][1] Introduce the new multiple resolution filter called the details [1.1] introduction singular point filter of the component engineering, and calculate matching between pictures correctly. The preliminary knowledge about an object is

unnecessary entirely. In each resolution, calculation of matching between pictures is calculated, while following the hierarchy of resolution. In that case, the hierarchy of resolution is followed from the coarse level in order to the minute level. A parameter required for calculation is thoroughly set up automatically by the dynamic calculation similar to human being's visual system. It is not necessary to specify the corresponding points between pictures with a help.

[0018] This premise art is applicable to generation of the smooth video from completely automatic morphing, object recognition, stereophotogrammetry, volume rendering, and few frames, etc., for example. When using for morphing, the given picture can be transformed automatically. When using for volume rendering, the interim picture between sections can be reconstructed correctly. It is the same, even when the shape of a long distance and a section is large and the distance between sections changes.

[0019][1.2] The multiplex resolution singular point filter concerning the hierarchy premise art of a singular point filter can save the luminosity and the position of each singular point which are moreover included in a picture, dropping the resolution of a picture. Width of a picture is set to N and height is set to M here. Below, since it is easy, it is assumed that it is $N=M=2^n$ (n is a natural number). 0 and section [N] **R is described to be I. The pixel of the picture in (i, j) is described to be $p_{(i,j)}$.

[0020] The hierarchy of multiplex resolution is introduced here. The hierarchized image group is generated by the multiple resolution filter. A multiple resolution filter generates another picture whose resolution is lower than the picture of even if two-dimensional search is performed to the picture of a basis, a singular point is detected and it extracts the detected singular point. Size of each picture in the m-th level is made into $2^m \times 2^m$ ($0 \leq m \leq n$) here. A singular point filter builds recursively the following four kinds of new hierarchy pictures towards falling from n.

[0021]

[Equation 1]

$$\begin{aligned} p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\ p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\ p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\ p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)})) \end{aligned}$$

(式1)

However, it is here,[Equation 2]

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式2})$$

It carries out. These four pictures are henceforth called a subsidiary image (sub image). If $\min_{x \leq t \leq x+1}$ and $\max_{x \leq t \leq x+1}$ are described to be alpha and beta, respectively, a subsidiary image can be described as follows, respectively.

[0022]

$P^{(m, 0)} = \text{alpha}(x) -- \text{alpha}(y)$ $p^{(m+1, 0)}$ $P^{(m, 1)} = \text{alpha}(x) -- \text{beta}(y)$ $p^{(m+1, 1)}$ $P^{(m, 2)} = \text{beta}(x) -- \text{alpha}(y)$ $p^{(m+1, 2)}$

$P^{(m, 3)} = \text{beta}(x)\text{beta}(y)$ $p^{(m+1, 3)}$, i.e., these, is considered to be a thing like a tensor product of alpha and beta. A subsidiary image supports a singular point, respectively. A singular point filter detects a singular point for every block which comprises 2x2 pixels about a picture of a basis so that clearly from these formulas. In that case, it searches for a point with a maximum pixel value or the minimum pixel value about 2 directions, i.e., length, and the side of each block. Although luminosity is adopted with premise art as a pixel value, various numerical values about a picture are employable. While a pixel from which a pixel which serves as a maximum pixel value about both 2 directions serves as the minimum pixel value about both a maximum point and 2 directions serves as a maximum pixel value about one side of a minimum point and 2 directions, a pixel which serves as the minimum pixel value about another side is detected as a saddle point.

[0023]A singular point filter drops resolution of a picture by representing a picture (here 4 pixels) of the block with a picture (here 1 pixel) of a singular point detected inside each block. If it carries out from a theoretical viewpoint of a singular point, $\text{alpha}(x)\text{alpha}(y)$ saves a minimum point, $\text{beta}(x)\text{beta}(y)$ saves a maximum point, and $\text{alpha}(x)\text{beta}(y)$ and $\text{beta}(x)\text{alpha}(y)$ save a saddle point.

[0024]First, singular point filtering is independently performed to a starting point (sauce) picture and a terminal point (destination) picture which should take matching, and a series of image groups, i.e., starting point hierarchy picture, and terminal point hierarchy pictures are generated, respectively. A starting point hierarchy picture and four kinds of terminal point hierarchy pictures are generated at a time corresponding to a kind of singular point, respectively.

[0025]Then, matching of a starting point hierarchy picture and a terminal point hierarchy picture can be taken in a series of resolution levels, and it is *****. Matching of a minimum point is first taken using $p^{(m, 0)}$. Next, based on the result, matching of a saddle point is taken using $p^{(m, 1)}$, and matching of other saddle points is taken using $p^{(m, 2)}$. And finally $p^{(m, 3)}$ is used and matching of a maximum point is taken.

[0026]Drawing 1 (c) and drawing 1 (d) show subsidiary image $p^{(5, 0)}$ of drawing 1 (a) and drawing 1 (b), respectively. Similarly, as for drawing 1 (e) and drawing 1 (f), $p^{(5, 1)}$, drawing 1 (g), and drawing 1 (h) show $p^{(5, 2)}$, respectively, as for $p^{(5, 2)}$, drawing 1 (i), and drawing 1 (j). According to the subsidiary image, matching of a characterizing portion of a picture becomes easy as these figures show. Eyes become clear by $p^{(5, 0)}$ first. It is because eyes are the minimum points of luminosity in a face. According to $p^{(5, 1)}$, a

mouth becomes clear. A mouth is because luminosity is low in a transverse direction. According to $p^{(5,2)}$, a vertical bar of both sides of a head becomes clear. The brightest point of an ear or a cheek becomes clear by $p^{(5,3)}$ at the last. It is because these are the maximum points of luminosity.

[0027] Since the feature of a picture can be extracted according to the singular point filter, a photographic subject reflected in a camera is discriminable by comparing the feature of some objects beforehand recorded as the feature of a picture photoed, for example with a camera.

[0028][1.3] Write a pixel of a position (i, j) of a calculation starting point picture of a map between pictures to be $p^{(n)}_{(i,j)}$, and, similarly $q^{(n)}_{(k,l)}$ describes a pixel of a position (k, l) of a terminal point picture. It is considered as i, j, k , and $I**I$. Energy (after-mentioned) of a map between pictures is defined. This energy is decided by a difference of luminosity of a pixel of a starting point picture, and luminosity of a pixel to which a terminal point picture corresponds, and smoothness of a map. Map $f^{(m,0)} : p^{(m,0)} \rightarrow q^{(m,0)}$ between $p^{(m,0)}$ which has the minimum energy in the beginning, and $q^{(m,0)}$ is calculated. Based on $f^{(m,0)}$, $p^{(m,1)}$ with the threshold energy and map $f^{(m,1)}$ between $q^{(m,1)}$ are calculated. This procedure continues until calculation of map $f^{(m,3)}$ between $p^{(m,3)}$ and $q^{(m,3)}$ is completed. Each map $f^{(m,i)}$ ($i = 0, 1, 2, \dots$) will be called a submap. For convenience of calculation of $f^{(m,i)}$, an order of i can be rearranged like a following formula. A reason to be rearranged is mentioned later.

[0029]

[Equation 3]

$$f^{(m,i)} : p^{(m,\sigma(i))} \rightarrow q^{(m,\sigma(i))} \quad (\text{式 3})$$

It is $\sigma(i) \in \{0, 1, 2, 3\}$ here.

[0030][1.3.1] When expressing matching between a bijection starting point picture and a terminal point picture by map, the map should fulfill bijection conditions between both images. It is because there are no superiority or inferiority on a concept in both images and a mutual pixel should be connected by surjection and a monomorphism. However, the map which should be built here unlike the usual case is the digital version of bijection. A pixel is specified by the lattice point in premise art.

[0031] The map to a terminal point subsidiary image (subsidiary image provided about the terminal point picture) from a starting point subsidiary image (subsidiary image provided about the starting point picture) is expressed by $f^{(m,s)} : I/2^{n-m} \times I/2^{n-m} \rightarrow I/2^{n-m} \times I/2^{n-m}$ ($s = 0, 1, \dots$). Here, $f^{(m,s)}(i, j) = (k, l)$ means that $p^{(m,s)}_{(i,j)}$ of a starting point picture is mapped by $q^{(m,s)}_{(k,l)}$ of a terminal point picture. Since it is easy, when $f(i, j) = (k, l)$ is realized, pixel $q_{(k,l)}$ is described to be $q_{f(i,j)}$.

[0032] Like a pixel (lattice point) treated with premise art, when data is discrete, a definition of bijection is important. Here, a definition is given as follows (all of i, i', j, j' , k ,

and I consider it as an integer). Each square regions first written by R in a flat surface of a starting point picture, [Equation 4]

$$P_{(i,j)}^{(m,s)} P_{(i+1,j)}^{(m,s)} P_{(i+1,j+1)}^{(m,s)} P_{(i,j+1)}^{(m,s)}$$

(式 4)

***** (i= 0, --, 2^m-1, j= 0, --, 2^m-1). The direction of each neighborhood (edge) of R is defined as follows here.

[0033]

[Equation 5]

$$\overrightarrow{P_{(i,j)}^{(m,s)} P_{(i+1,j)}^{(m,s)}}, \overrightarrow{P_{(i,j)}^{(m,s)} P_{(i+1,j+1)}^{(m,s)}}, \overrightarrow{P_{(i+1,j)}^{(m,s)} P_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{P_{(i,j+1)}^{(m,s)} P_{(i,j)}^{(m,s)}}$$

(式 5)

This square must be mapped by the map f by the quadrilateral in a terminal point image plane. The quadrilateral shown by $f^{(m,s)}(R)$, [Equation 6]

$$q_{f(i,j)}^{(m,s)} q_{f(i+1,j)}^{(m,s)} q_{f(i+1,j+1)}^{(m,s)} q_{f(i,j+1)}^{(m,s)}$$

(式 6)

It is necessary to fulfill ** and the following bijection conditions.

[0034] 1. Edge of quadrilateral $f^{(m,s)}(R)$ does not cross mutually.

2. The direction of edge of $f^{(m,s)}(R)$ is equal to them of R (in the case of drawing 2 clockwise rotation).

3. Allow a contraction map (RITORA cushion: retractions) as relaxation conditions.

[0035] Unless some relaxation conditions are established, it is because a map which fulfills bijection conditions thoroughly has only a unit map. Here, the length of one edge of $f^{(m,s)}(R)$ may become a triangle 0 (R), i.e., $f^{(m,s)}$. However, don't become a figure, i.e., one point, that area is set to 0, or one line segment. When drawing 2 (R) is a quadrilateral of a basis, although drawing 2 (A) and drawing 2 (D) fulfill bijection conditions, drawing 2 (B), drawing 2 (C), and drawing 2 (E) do not fill them.

[0036] In a actual implementation, the following conditions may be imposed further that it should guarantee easily that a map is surjection. That is, as for each pixel on a boundary of a starting point picture, a pixel which occupies the same position in a terminal point picture ****. That is, it is $f(i, j) = (i, j)$ (however, on four lines of $i=0$, $i=2^m-1$, $j=0$, and $j=2^m-1$). This condition is also called the following "additional condition."

[0037] [1.3.2] Define energy of the cost map f about luminosity of an energy [1.3.2.1] pixel of a map. It is the purpose that energy looks for a map which becomes the minimum. Energy is mainly decided by a difference of luminosity of a pixel of a starting point picture, and luminosity of a pixel of a terminal point picture corresponding to it. That is, energy $C_{(i,j)}^{(m,s)}$ in a point (i, j) of map $f^{(m,s)}$ becomes settled with a following formula.

[0038]

[Equation 7]

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_{f(i,j)}^{(m,s)})|^2 \quad (\text{式 7})$$

Here, $V(p_{(i,j)}^{(m,s)})$ and $V(q_{f(i,j)}^{(m,s)})$ are the luminosity of pixel $p_{(i,j)}^{(m,s)}$ and $q_{f(i,j)}^{(m,s)}$, respectively. Total energy $C_f^{(m,s)}$ of f is one valuation plan by which matching is evaluated, and the sum total of $C_{(i,j)}^{(m,s)}$ shown below can define it.

[0039]

[Equation 8]

$$C_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} C_{(i,j)}^{(m,s)} \quad (\text{式 8})$$

[1.3.2.2] In order to obtain the cost smooth map about the position of the pixel for a smooth map, introduce another energy D_f about a map. This energy is decided by the position of $p_{(i,j)}^{(m,s)}$ and $q_{f(i,j)}^{(m,s)}$ regardless of the luminosity of a pixel ($i=0, \dots, 2^m-1, j=0, \dots, 2^m-1$). Energy $D_{[of map f in a point (i,j)]}^{(m,s)}$ is defined by the following formula.

[0040]

[Equation 9]

$$D_{(i,j)}^{(m,s)} = \eta E_0^{(m,s)} + E_1^{(m,s)} \quad (\text{式 9})$$

however -- coefficient-parameters eta is the zero or more real numbers -- moreover [Equation 10]

$$E_0^{(m,s)} = \|(i,j) - f^{(m,s)}(i,j)\|^2 \quad (\text{式 10})$$

[Equation 11]

$$E_1^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2 / 4 \quad (\text{式 11})$$

It carries out. It is here, [Equation 12]

$$\|(x,y)\| = \sqrt{x^2 + y^2} \quad (\text{式 12})$$

It comes out, and it is and $f(i',j')$ decides on 0 to $i' < 0$ and $j' < 0$. E_0 is decided by distance of (i,j) and $f(i,j)$. E_0 prevents ***ing a pixel to the pixel left too much. However, E_0 is replaced with behind another energy function. E_1 guarantees the smoothness of a map. E_1 expresses the distance between displacement and

displacement of an adjacent point of $p_{(i,j)}$. Based on the above consideration, energy D_f , which evaluates matching and which is another valuation plan becomes settled with a following formula.

[0041]

[Equation 13]

$$D_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} D_{(i,j)}^{(m,s)} \quad (\text{式 1 3})$$

[1.3.2.3] The overall evaluation type concerning integration of the total calories of a total-calories map of a map, i.e., two or more valuation plans, is defined by $\lambda C_f^{(m,s)} + D_f^{(m,s)}$. Coefficient-parameters lambda is the zero or more real numbers here. The purpose is for an overall evaluation type to detect the state of taking an extremum, i.e., find out the map which gives the threshold energy shown with a following formula.

[0042]

[Equation 14]

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \quad (\text{式 1 4})$$

In the case of $\lambda = 0$ and $\eta = 0$, it should be cautious of a map turning into a unit map (that is, it becomes $f^{(m,s)}(i,j) = (i,j)$ to all the $i = 0, \dots, 2^m-1$ and $j = 0, \dots, 2^m-1$). Since this premise art estimates the case of $\lambda = 0$ and $\eta = 0$ first like the after-mentioned, a map can be gradually transformed from a unit map. Supposing it changes the position of λ of an overall evaluation type temporarily and defines it as $C_f^{(m,s)} + \lambda D_f^{(m,s)}$, The pixels where an overall evaluation type becomes only $C_f^{(m,s)}$, and which in the case of $\lambda = 0$ and $\eta = 0$ originally do not have relation in any way are matched only by saying [that luminosity is only near], and a map will become meaningless. Even if it transforms the map based on such a meaningless map, a meaning is not made at all. For this reason, it is a unit map at the start time of evaluation, and is considered in how to give coefficient parameters so that it may be chosen as the best map.

[0043] An optical flow as well as this premise art takes into consideration a difference and smoothness of luminosity of a pixel. However, an optical flow cannot be used for conversion of a picture. It is for taking only a local motion of an object into consideration. A global correspondence relation is detectable by using a singular point filter concerning premise art.

[0044][1.3.3] Give the determination threshold energy of a map by introduction of multiplex resolution, and calculate map f_{\min} which satisfies bijection conditions using a hierarchy of multiplex resolution. In each resolution level, a map between a starting point subsidiary image and a terminal point subsidiary image is calculated. It starts from the top (coarsest level) of a hierarchy of resolution, and a map of each resolution

level is determined, taking a map of other levels into consideration. The number of candidates of a map in each level is restricted by using a map of a coarser level more highly that is,. More specifically, a map called for in a level coarser than it by one is imposed as a kind of constraint when determining a map in a certain level.

[0045]first[Equation 15]

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式 1 5})$$

***** -- it comes suddenly and $p^{(m-1, s)}_{(i', j')}$ and $q^{(m-1, s)}_{(i', j')}$ will be called parent of $p^{(m, s)}_{(i, j)}$ and $q^{(m, s)}_{(i, j)}$, respectively. [x] It is the maximum integer which does not exceed x. $p^{(m, s)}_{(i, j)}$ and $q^{(m, s)}_{(i, j)}$ are called child of $p^{(m-1, s)}_{(i', j')}$ and $q^{(m-1, s)}_{(i', j')}$, respectively. The function parent (i, j) is defined by following formula.

[0046]

[Equation 16]

$$\text{parent}(i, j) = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式 1 6})$$

Map $f^{(m, s)}$ between $p^{(m, s)}_{(i, j)}$ and $q^{(m, s)}_{(k, l)}$ is determined by finding what performed energy calculation and became the minimum. By using $f^{(m-1, s)}$ ($m=1, 2, \dots, n$), the value of $f^{(m, s)}_{(i, j)} = (k, l)$ is determined as follows. First, $q^{(m, s)}_{(k, l)}$ imposes the conditions that it must be in the inside of the following quadrilateral, and narrows down what has high actuality among the maps which fulfill bijection conditions.

[0047]

[Equation 17]

$$q^{(m, s)}_{g^{(m, s)}(i-1, j-1)} q^{(m, s)}_{g^{(m, s)}(i-1, j+1)} q^{(m, s)}_{g^{(m, s)}(i+1, j+1)} q^{(m, s)}_{g^{(m, s)}(i+1, j-1)}$$

$$(\text{式 1 7})$$

However, it is here.[Equation 18]

$$g^{(m, s)}_{(i, j)} = f^{(m-1, s)}(\text{parent}(i, j)) + f^{(m-1, s)}(\text{parent}(i, j) + (1, 1))$$

$$(\text{式 1 8})$$

It comes out. In this way, the defined quadrilateral is made to call it below the inheritance (inherited) quadrilateral of $p^{(m, s)}_{(i, j)}$. It asks for the pixel which makes energy the minimum in the inside of an inheritance quadrilateral.

[0048]Drawing 3 shows the above procedure. said -- a figure -- setting -- the starting point -- a picture -- A -- B -- C -- D -- a pixel -- the -- m - one -- a level -- setting -- respectively -- a terminal point -- a picture -- A -- ' -- B -- ' -- C -- ' -- D -- ' -- ****(ing) -- having . Pixel $p^{(m, s)}_{(i, j)}$ must be ****(ed) to pixel $q^{(m, s)}_{f(m)(i, j)}$ which exists in the inside of inheritance quadrilateral A'B'C'D'. The mediation to the map of the m-th level from a map of the m-1st levels is made by the above

consideration.

[0049] Energy E_0 defined previously is transposed to a following formula, in order to calculate submap $f^{(m, 0)}$ in the m-th level.

[0050]

[Equation 19]

$$E_{0(i,j)} = \|f^{(m,0)}(i,j) - g^{(m)}(i,j)\|^2 \quad (\text{式 19})$$

A following formula is used in order to calculate submap $f^{(m, s)}$.

[0051]

[Equation 20]

$$E_{0(i,j)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2 \quad (1 \leq i) \quad (\text{式 20})$$

In this way, the map which maintains the energy of all the submaps at a low value is obtained. The submap corresponding to a different singular point is associated within an identical level so that the similarity of submaps may become high by the formula 20. The formula 19 shows distance with the position of the point that $f^{(m, s)}(i, j)$, and (i, j) at the time of considering it a part of pixel of the m-1st levels should be projected.

[0052] The following measures are taken when a pixel which fulfills bijection conditions does not exist in an inside of inheritance quadrilateral A'B'C'D' temporarily. First, distance from a boundary line of A'B'C'D' investigates a pixel which is L (it is L=1 in the beginning). Among those, if that from which energy becomes the minimum fulfills bijection conditions, this will be chosen as a value of $f^{(m, s)}(i, j)$. L is enlarged until such a point is discovered or L reaches $L^{(m)}$ max of the maximum. $L^{(m)}$ max is immobilization to each level m. When such a point is not discovered at all, map that the 3rd condition of bijection is disregarded temporarily and area of a quadrilateral of a converting destination becomes zero is also accepted, and $f^{(m, s)}(i, j)$ is determined. When a point of still fulfilling conditions is not found, the 1st and 2nd conditions of bijection are removed next.

[0053] Avoiding that a map is influenced by details of a picture, an approximation using multiplex resolution is indispensable in order to determine a global correspondence relation between pictures. If an approximation by multiplex resolution is not used, it is impossible to find out a correspondence relation between pixels with a far distance. In that case, size of a picture must be limited to a very small thing, and only a small picture of change can be treated. A correspondence relation between such pixels is made hard to find, in order to usually require smoothness of a map. It is because energy of a map to a pixel with distance from a pixel is high. According to the approximation using multiplex resolution, a suitable correspondence relation between such pixels can be found out. Those distance is because it is small in an upper level

(coarse level) of a hierarchy of resolution.

[0054][1.4] There is difficulty of one PAREMETA adjustment of the main faults of matching art of automatic determination existing of optimal PAREMETA value. When the most, it is very difficult for adjustment of a parameter to be performed by help work and to choose optimal value. According to a method concerning premise art, automatic determination of the optimal parameter value can be carried out thoroughly.

[0055]A system concerning premise art contains 2 PAREMETA, and lambda and eta. Speaking directly, lambda's being the dignity of a difference of luminosity of a pixel and eta's showing the rigidity of a map. An initial value is 0, it fixes to eta= 0 first and a value of these parameters makes lambda increase from 0 gradually. When making a value of an overall evaluation type (formula 14) into the minimum moreover, enlarging a value of lambda, generally a value of $C^{(m, s)}_f$, about each submap becomes small. This means that 2 pictures must match more fundamentally. However, if lambda exceeds an optimum value, the following phenomena will occur.

[0056]1. The pixels which should not correspond essentially are accidentally matched only by saying [that luminosity is only near].

2. As a result, a correspondence relation of pixels becomes amusing and a map begins to collapse.

[0057]3. As a result, in the formula 14, $D^{(m, s)}_f$ tends to increase rapidly.

4. As a result, since a value of the formula 14 tends to increase rapidly, $f^{(m, s)}$ changes so that a rapid increase in $D^{(m, s)}_f$ may be controlled, and $C^{(m, s)}_f$ increases.

Therefore, maintaining the state of taking the minimum making lambda increase, $C^{(m, s)}_f$ detects a threshold changed to an increase from reduction, and the formula 14 makes the lambda an optimum value in eta= 0. Next, eta is increased little by little, an action of $C^{(m, s)}_f$ is inspected, and automatic determination of eta is carried out by a below-mentioned method. lambda is also decided corresponding to the eta.

[0058]This method resembles operation of a focal mechanism human being's visual system. In human being's visual system, matching of a picture of right-and-left both eyes is taken, moving one eye. The eye is fixed when an object can recognize clearly.

[0059][1.4.1] It is increased from 0 by the dynamic determination lambda of lambda by predetermined unit width, and a submap is evaluated whenever a value of lambda changes. Total calories are defined by lambda $C^{(m, s)}_f + D^{(m, s)}_f$ like the formula 14. $D^{(m, s)}_f$ of the formula 9 expresses smoothness, and, in a unit map, becomes the minimum theoretically, and E_0 and E_1 also increases it, so that a map is distorted. Since E_1 is an integer, the minimum unit width of $D^{(m, s)}_f$ is 1. For this reason, if change (decrement) of the present lambda $C^{(m, s)}_{(i, j)}$ is not one or more, total calories cannot be reduced by changing a map. Because, it is for total calories not decreasing, unless lambda $C^{(m, s)}_{(i, j)}$ decreases or more by one, since $D^{(m, s)}_f$ increases one or more with change of a map.

[0060]With a basis of this condition, and an increase in lambda, when normal, it is shown that $C^{(m, s)}_{(i, j)}$ decreases. A histogram of $C^{(m, s)}_{(i, j)}$ is described to be $h(l)$. $h(l)$ is

the number of pixels whose energy $C_{(i,j)}^{(m,s)}$ are I^2 . Since $\lambda \rightarrow 1$ is realized, a case of $I^2=1/\lambda$ is considered, for example. When λ carries out fine amount change from λ_1 to λ_2 , [Equation 21]

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{l=\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式 2 1)

It comes out and A pixels shown,[Equation 22]

$$C_f^{(m,s)} - I^2 = C_f^{(m,s)} - \frac{1}{\lambda}$$

(式 2 2)

It changes [rather than] to a stable state with ** energy. Here, temporarily, if all of the energy of these pixels become zero, it approximates. The value of $C_f^{(m,s)}$, this formula, [Equation 23]

$$\partial C_f^{(m,s)} = -\frac{A}{\lambda}$$

(式 2 3)

Carrying out *** change is shown and it is the result,[Equation 24]

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}}$$

(式 2 4)

It *****. $h(l)$ Since it is >0 , $C_f^{(m,s)}$ usually decreases. However, when λ tends to exceed an optimum value, an above-mentioned phenomenon, i.e., the increase in $C_f^{(m,s)}$, occurs. The optimum value of λ is determined by detecting this phenomenon.
[0061]When making H ($h > 0$) and k into a constant,[Equation 25]

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}}$$

(式 2 5)

It is if it assumes, [Equation 26]

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}}$$

(式 2 6)

*****. If it is $k!=-3$ at this time, [Equation 27]

$$C_f^{(m,s)} = C + \frac{H}{(3/2+k/2)\lambda^{3/2+k/2}}$$

(式 2 7)

It becomes. This is a general formula of $C_f^{(m,s)}$ (C is a constant).

[0062]When detecting the optimum value of λ , safety may be seen further and the number of the pixels which break bijection conditions may be inspected. When determining the map of each pixel here, the probability of breaking bijection conditions is assumed to be p_0 . In this case [Equation 28]

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式 28})$$

Since it is *****ing), the number of pixels which break bijection conditions increases at a rate of a following formula.

[0063]

[Equation 29]

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}}$$

(式 29)

It follows, [Equation 30]

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1 \quad (\text{式 30})$$

It is *****. the time of assuming $h(l) = H l^k$ temporarily -- for example [Equation 31]

$$B_0 \lambda^{3/2+k/2} = p_0 H \quad (\text{式 31})$$

It becomes *****. However, if lambda exceeds an optimum value, the upper value will increase quickly. This phenomenon can be detected, it can be inspected whether the value of $B_0 \lambda^{3/2+k/2}/2^m$ exceeds abnormal value $B_{0\text{thres}}$, and the optimum value of lambda can be determined. Rate-of-increase B_1 of a pixel which breaks the 3rd condition of bijection is checked by similarly inspecting whether the value of $B_1 \lambda^{3/2+k/2}/2^m$ exceeds abnormal value $B_{1\text{thres}}$. The reason for introducing factor 2^m is mentioned later. This system is not sensitive to these two thresholds. By observation of energy $C^{(m, s)}$, these thresholds can be used in order to detect the excessive distortion of the map which it failed to detect.

[0064] In an experiment, when calculating submap $f^{(m, s)}$, if lambda exceeded 0.1, calculation of $f^{(m, s)}$ was stopped and shifted to calculation of $f^{(m, s+1)}$. At the time of $\lambda > 0.1$, it is because [in luminosity 255 level of a pixel] a difference in "3" influenced calculation of a submap only, and it is because it was difficult to obtain a right result at the time of $\lambda > 0.1$.

[0065][1.4.2] Histogram $h(l)$

It does not depend for an inspection of $C^{(m, s)}$ on the histogram $h(l)$. In the case of an inspection of bijection and its 3rd condition, influence can be received in $h(l)$. When it actually ($\lambda, C^{(m, s)}$) plots, k is usually in the one neighborhood. In an experiment, $B_0 \lambda^2$ and $B_1 \lambda^2$ were inspected using $k=1$. If a true value of k is less than one, $B_0 \lambda^2$ and $B_1 \lambda^2$ will not become a constant but will increase gradually according to factor $\lambda^{(1-k)/2}$. If $h(l)$ is a constant, factors are $\lambda^{1/2}$, for example. However, such a difference is absorbable by setting up threshold $B_{0\text{thres}}$ correctly.

[0066] A center assumes a starting point picture here to be (x_0, y_0) , and a circular object of the radius r like a following formula.

[Equation 32]

$$p(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_0)^2 + (j-y_0)^2}) & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式 3 2})$$

On the other hand, a terminal point picture presupposes that a center (x_1, y_1) and a radius are the objects of r like a following formula.

[Equation 33]

$$q(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式 3 3})$$

$c(x)$ presupposes here that it is a form of $c(x)=x^k$. In the histogram $h(l)$, a center (x_0, y_0) , and (x_1, y_1) serve as a form of a following formula, when far enough.

[Equation 34]

$$h(l) \propto r l^k \quad (k \neq 0) \quad (\text{式 3 4})$$

A picture shows an object with a clear boundary line embedded for a background at the time of $k=1$. This object becomes bright as a center is dark and goes to the circumference. A picture expresses an object with an ambiguous boundary line at the time of $k=-1$. This object becomes dark as a center is the brightest and goes to the circumference. Even if it thinks that there is a general object in the middle of these 2 types of object, generality is not very lost. Therefore, k can cover a case where it is the most as $-1 \leq k \leq 1$, and it is secured that the formula 27 is generally a decreasing function.

[0067] As shown in the formula 34, it should be cautious of r being influenced by resolution of a picture, i.e., r is proportional to $2 m$. For this reason, $2 m$ of factors were introduced in [1.4.1].

[0068][1.4.3] The automatic determination of the dynamic determination parameter η of η can be carried out in a similar way. It is first referred to as $\eta=0$, and final map $f^{(n)}$ and energy $C^{(n)}_f$ in the finest resolution are calculated. Continue, only a certain value $\Delta\eta$ makes η increase, and last map $f^{(n)}$ and energy $C^{(n)}_f$ in the again finest resolution is recalculated. This process is continued until an optimum value can be found. η shows the rigidity of a map. It is because it is the dignity of a following formula.

[0069]

[Equation 35]

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式 3 5)

When eta is 0, $D^{(n)}_f$ will be determined regardless of the last submap, will be elastically transformed by the present submap, and will be distorted too much. On the other hand, as for $D^{(n)}_f$, eta is decided by the last submap nearly thoroughly at the time of a very big value. At this time, a submap has dramatically high rigidity, and a pixel is projected on the same place. As a result, a map turns into a unit map. When the value of eta increases gradually from 0, $C^{(n)}_f$ decreases gradually like the after-mentioned. However, energy begins to increase as it is shown in drawing 4, when the value of eta exceeds an optimum value. The X-axis of the figure is eta and a Y-axis is C_f .

[0070] The value of optimal eta that makes $C^{(n)}_f$ the minimum by this method can be obtained. However, as a result of various elements' influencing calculation compared with the case of lambda, $C^{(n)}_f$ changes, swinging small. In eta, it is because all the submaps are recalculated but as a submap is recalculated once, whenever an input carries out fine amount change of the case of lambda. For this reason, it cannot be judged immediately whether the value of obtained $C^{(n)}_f$ is the minimum. If the candidate of the minimum is found, it is necessary to look for the true minimum by setting up the still finer section.

[0071][1.5] When determining a correspondence relation between super sampling pixels, in order to increase flexibility, a range of $f^{(m,s)}$ is extensible to RxR (R is a set of the real number). In this case, luminosity of a pixel of a terminal point picture is interpolated and it is a nonintegral point, [Equation 36]

$$V(q_{f^{(m,s)}(i,j)}^{(m,s)}) \quad (\text{式36})$$

$f^{(m,s)}$ with the luminosity which can be boiled and set is provided. That is, super sampling is performed. In an experiment, $f^{(m,s)}$ is allowed to take an integer and a half-integral value, [Equation 37]

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)}) \quad (\text{式37})$$

**,[Equation 38]

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2 \quad (\text{式38})$$

Be alike was given.

[0072][1.6] When a normalization starting point picture and a terminal point picture of luminosity of each picture contain an extremely different object, it is hard to use luminosity of the original pixel for calculation of a map as it is. [of a pixel] since the difference of luminosity is large, energy $C^{(m,s)}_f$ about luminosity becomes large too much — right evaluation — butter fish — ** — a sake — it is .

[0073]For example, a case where matching of people's face and a face of a cat is taken is considered. A face of a cat is covered with hair and a very bright pixel and a very dark pixel are mixed. In this case, in order to calculate a submap between 2 faces, a subsidiary image is normalized first. That is, luminosity of the darkest pixel is set as 0, the brightest it is set as 255, and it asks for luminosity of other pixels by linear interpolation.

[0074][1.7] Use how calculation advances linearly according to a scan of an implementation starting point picture inductive. A value of $f^{(m, s)}$ is determined about pixel [of a left end on introduction and No. 1] $(i, j) = (0, 0)$. Next, a value of each $f^{(m, s)}$ (i, j) is determined, increasing i every [1]. When a value of i reaches width of a picture, a value of j is increased one and i is returned to 0. Henceforth, $f^{(m, s)}(i, j)$ is determined with a scan of a starting point picture. If correspondence of a pixel is decided about all the points, one map $f^{(m, s)}$ will be decided.

[0075]If corresponding-points $q_{R(i, j)}$ is decided about certain $p_{(i, j)}$, corresponding-points $q_{R([p(i, j+1)](i, j+1))}$ will be decided next. Under the present circumstances, a position of $q_{R(i, j+1)}$ is restricted by position of $q_{R(i, j)}$ in order to fulfill bijection conditions. Therefore, in a point that corresponding points are decided previously, in this system, a priority becomes high. An excessive deviation will be added to the last map called for if the state where priority is always $(0, 0)$ given most continues. With this premise art, in order to avoid this state, $f^{(m, s)}$ is decided by the following methods.

[0076]When it is 0 first (second mod 4), it decides making $(0, 0)$ into the starting point and increasing i and j gradually. When (second mod 4) is 1, a right end point of the top line is made into the starting point, and i is decided, making reduction and j increase. When (second mod 4) is 2, a right end point of a lowermost row is made into the starting point, and it decides, decreasing i and j . When (smod 4) is 3, a left end point of a lowermost row is made into the starting point, and i is decided, decreasing an increase and j . Since the concept s of a submap, i.e., a parameter, did not exist in the n -th level with the finest resolution, even if it was $s=0$ and $s=2$, 2 directions were calculated continuously.

[0077]In a actual implementation, a value of $f^{(m, s)}(i, j)$ ($m=0, \dots, n$) which fulfills bijection conditions as much as possible out of a candidate (k, l) was chosen by giving a penalty to a candidate who breaks bijection conditions. phi is hung on the energy D of a candidate who breaks the 3rd condition (k, l) , and, on the other hand, psi is hung on a candidate who breaks the 1st or 2nd condition. phi=2 and psi=100000 were used this time.

[0078]The following tests were done when determining $=f^{(m, s)}(i, j)$ as a procedure actual for a check of the above-mentioned bijection conditions (k, l) . That is, it is confirmed whether z ingredient of an outer product of a following formula becomes zero or more to each lattice point (k, l) contained in an inheritance quadrilateral of $f^{(m, s)}(i, j)$.

[0079]

[Equation 39]

$$W = \vec{A} \times \vec{B}$$

(式 3 9)

However, it is here,[Equation 40]

$$\vec{A} = \overrightarrow{q_{f^{(m,s)}(i,j-1)}^{(m,s)} q_{f^{(m,s)}(i+1,j-1)}^{(m,s)}}$$

(式 4 0)

[Equation 41]

$$\vec{B} = \overrightarrow{q_{f^{(m,s)}(i,j-1)}^{(m,s)} q_{(k,l)}^{(m,s)}}$$

(式 4 1)

Come out and it is (a vector is made into a three-dimensional vector here, and the z-axis is defined in a rectangular right-handed coordinate system). If W is negative, it will be made whether a penalty is given and is made and not to choose by hanging psi on $D_{(k,l)}^{(m,s)}$ about the candidate.

[0080] Drawing 5 (a) and drawing 5 (b) show the reason for inspecting this condition.

Drawing 5 (a) expresses a candidate without a penalty, and the candidate to whom drawing 5 (b) has a penalty, respectively. When determining map $f^{(m,s)}(i, j+1)$ to an adjacent pixel $(i, j+1)$, if z ingredient of W is negative, the pixel which satisfies bijection conditions on a starting point image plane does not exist. Because, it is for $q_{(k,l)}^{(m,s)}$ exceeding the boundary line of the adjoining quadrilateral.

[0081][1.7.1] In an order implementation of a submap, when the number of resolution levels was even, $\sigma(0) = 3$, $\sigma(1) = 2$, $\sigma(2) = 1$, $\sigma(3) = 0$, and $\sigma(4) = 3$ were used using $\sigma(0) = 0$, $\sigma(1) = 1$, $\sigma(2) = 2$, $\sigma(3) = 3$, and $\sigma(4) = 0$ at the time of odd number. By this, a submap was shuffled moderately. Originally a submap is four kinds and s is either 0-3. However, processing which is actually equivalent to $s=4$ was performed. The reason is mentioned later.

[0082][1.8] After a map between an interpolation calculation starting point picture and a terminal point picture is determined, luminosity which is a pixel which corresponds and suits is interpolated. Try linear interpolation was used in an experiment. Square p

in a starting point image plane. (i, j) It is assumed that $P_{(i+1,j)} P_{(i,j+1)} P_{(i+1,j+1)}$ is projected on quadrilateral $q_{f(i,j)} q_{f(i+1,j)} q_{f(i,j+1)} q_{f(i+1,j+1)}$ on a terminal point image plane. Since it is easy, distance between pictures is set to 1. The pixel r (x, y, t) of an intermediate image whose distance from a starting point image plane is t ($0 \leq t \leq 1$) ($0 \leq x \leq N-1$, $0 \leq y \leq M-1$) is called for in the following ways. A position (however, $x, y, t**R$) of the pixel r (x, y, t) is first searched for with a following formula.

[0083]

[Equation 42]

$$\begin{aligned}
 (x, y) = & (1 - dx)(1 - dy)(1 - t)(i, j) + (1 - dx)(1 - dy)t f(i, j) \\
 & + dx(1 - dy)(1 - t)(i + 1, j) + dx(1 - dy)t f(i + 1, j) \\
 & + (1 - dx)dy(1 - t)(i, j + 1) + (1 - dx)dyt f(i, j + 1) \\
 & + dxdy(1 - t)(i + 1, j + 1) + dxdytf(i + 1, j + 1)
 \end{aligned}$$

(式 4 2)

The luminosity of the pixel in $r(x, y, t)$ is continuously determined using the following formula.

[0084]

[Equation 43]

$$\begin{aligned}
 V(r(x, y, t)) = & (1 - dx)(1 - dy)(1 - t)V(p_{(i,j)}) + (1 - dx)(1 - dy)tV(q_{f(i,j)}) \\
 & + dx(1 - dy)(1 - t)V(p_{(i+1,j)}) + dx(1 - dy)tV(q_{f(i+1,j)}) \\
 & + (1 - dx)dy(1 - t)V(p_{(i,j+1)}) + (1 - dx)dytV(q_{f(i,j+1)}) \\
 & + dxdy(1 - t)V(p_{(i+1,j+1)}) + dxdytfV(q_{f(i+1,j+1)})
 \end{aligned}$$

(式 4 3)

dx and dy are parameters and change from 0 to 1 here.

[0085][1.9] The determination of the map in case a constraint does not exist at all until now [map] when a constraint is imposed was described. However, when correspondence relations are beforehand specified between the specific pixels of a starting point picture and a terminal point picture, a map can be determined after making this into a constraint.

[0086]By the rough map which moves the specific pixel of a starting point picture to the specific pixel of a terminal point picture first, a fundamental idea transforms a starting point picture roughly, and calculates the map f correctly after an appropriate time.

[0087]It opts for a rough map which projects a specific pixel of a starting point picture on a specific pixel of a terminal point picture first, and projects other pixels of a starting point picture on a suitable position. That is, a pixel near a specific pixel is the map which is projected near the place where the specific pixel is projected. A rough map of the m -th level is described to be $F^{(m)}$ here.

[0088]It opts for the rough map F in the following ways. First, a map is specified about some pixels. It is a pixel of n_s individual about a starting point picture, [Equation 44]
 $P(i_0, j_0), P(i_1, j_1), \dots, P(i_{n_s-1}, j_{n_s-1})$

(式 4 4)

The following values are decided when specified.

[Equation 45]

$$\begin{aligned} F^{(n)}(i_0, j_0) &= (k_0, l_0), \\ F^{(n)}(i_1, j_0) &= (k_1, l_1), \dots, \\ F^{(n)}(i_{n_s-1}, j_{n_s-1}) &= (k_{n_s-1}, l_{n_s-1}) \end{aligned}$$

(式 4 5)

The amount of displacement of other pixels of a starting point picture is an average which makes weighting displacement of $p_{(i_h, j_h)}$ ($h = 0, \dots, n_s - 1$), and is called for. That is, pixel $p_{(i, j)}$ is projected on the following pixels of a terminal point picture.

[0089]

[Equation 46]

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{h=n_s-1} (k_h - i_h, l_h - j_h) weight_h(i, j)}{2^{n-m}}$$

(式 4 6)

However, it is here,[Equation 47]

$$weight_h(i, j) = \frac{1 / \|(i_h - i, j_h - j)\|^2}{total\ weight(i, j)}$$

(式 4 7)

[Equation 48]

$$total\ weight(i, j) = \sum_{h=0}^{h=n_s-1} 1 / \|(i_h - i, j_h - j)\|^2$$

(式 4 8)

It carries out.

[0090]It continues, and energy $D^{(m, s)}_{(i, j)}$ of the map f is changed so that the candidate map f near $F^{(m)}$ may have less energy. Correctly, $D^{(m, s)}_{(i, j)}$ is, [Equation 49]

$$D^{(m, s)}_{(i, j)} = E^{(m, s)}_{0(i, j)} + \eta E^{(m, s)}_{1(i, j)} + \kappa E^{(m, s)}_{2(i, j)}$$

(式 4 9)

It comes out. It corrects,[Equation 50]

$$E^{(m, s)}_{2(i, j)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i, j) - f^{(m, s)}(i, j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i, j) - f^{(m, s)}(i, j)\|^2, & \text{otherwise} \end{cases}$$

(式 5 0)

It comes out, and it is and is referred to as kappa and rho>=0. f is thoroughly determined as the last by the automatic calculation process of the above-mentioned map.

[0091]Here, in $f^{(m, s)}(i, j)$, when close to $F^{(m)}(i, j)$ enough that is, those distance,

[Equation 51]

$$\left[\frac{\rho^2}{2^{2(n-m)}} \right] \quad (\text{式 51})$$

When it is less than, it should be cautious of $E_2^{(m, s)}(i, j)$ being set to 0. The reason defined such is because he would like to decide the value automatically so that it may settle in a suitable position in a terminal point picture, as long as each $f^{(m, s)}(i, j)$ is close to $F^{(m)}(i, j)$ enough. It is automatically mapped so that it is not necessary to specify an exact correspondence relation in detail and a starting point picture may match a terminal point picture for this reason.

[0092][2] Explain a flow of processing by each component engineering of concrete procedure [1]. Drawing 6 is a flow chart which shows the whole premise technical procedure. As shown in the figure, processing which used a multiplex resolution singular point filter first is performed (S1), and matching of a starting point picture and a terminal point picture is taken continuously (S2). However, S2 may process image recognition etc. based on the feature of a picture acquired by S1 rather than is indispensable.

[0093]Drawing 7 is a flow chart which shows details of S1 of drawing 6. Here, it is premised on taking matching of a starting point picture and a terminal point picture by S2. Therefore, first, with a singular point filter, a starting point picture is hierarchized (S10) and a series of starting point hierarchy pictures are acquired. A terminal point picture is continuously hierarchized in a similar way (S11), and a series of terminal point hierarchy pictures are acquired. However, an order of S10 and S11 is arbitrary, and can also generate a starting point hierarchy picture and a terminal point hierarchy picture in parallel.

[0094]Drawing 8 is a flow chart which shows details of S10 of drawing 7. Size of a starting point picture of a basis is made into $2^n \times 2^n$. Since a starting point hierarchy picture is made sequentially from the one where resolution is finer, it sets to n the parameter m which shows a resolution level which is the target of processing (S100). Continuously Picture $p^{(m, 0)}$ of the m-th level, $p^{(m, 1)}, p^{(m, 2)}$, A singular point is detected from $p^{(m, 3)}$ using a singular point filter (S101), and picture $p^{(m-1, 0)}$ of the m-1st levels, $p^{(m-1, 1)}, p^{(m-1, 2)}$, and $p^{(m-1, 3)}$ are generated, respectively (S102). Here, since it is $m=n$, it is $p^{(m, 0)}=p^{(m, 1)}=p^{(m, 2)}=p^{(m, 3)}=p^{(n)}$, and four kinds of subsidiary images are generated from one starting point picture.

[0095]Drawing 9 shows correspondence relation between a part of picture of the m-th level, and a part of picture of the m-1st levels. A numerical value of the figure shows luminosity of each pixel. $p^{(m, s)}$ of the figure symbolizes four pictures of $p^{(m, 0)} - p^{(m, 3)}$, and in generating $p^{(m-1, 0)}$, it thinks that $p^{(m, s)}$ is $p^{(m, 0)}$. Under a rule shown by [1.2], $p^{(m-1, 0)}$ about a block as which luminosity was filled in for example, with the figure. $p^{(m-1, 1)}$ replaces "8" and $p^{(m-1, 2)}$ among 4 pixels "3" contained there by one pixel which

acquired "10" for "6" and $p^{(m-1,3)}$, respectively, and acquired this block, respectively. Therefore, size of a subsidiary image of the m-1st levels becomes $2^{m-1} \times 2^{m-1}$.

[0096]The decrement of the m is carried out continuously (S103 of drawing 8), and it checks that m is not negative (S104), it returns to S101, and then a subsidiary image with coarse resolution is generated. As a result of this repetition processing, when a subsidiary image of m= 0, i.e., the 0th level, is generated, S10 is completed. Size of a subsidiary image of the 0th level is 1x1.

[0097]Drawing 10 has illustrated a starting point hierarchy picture generated by S10 about a case of n= 3. Only the first starting point picture is common to four series, and a subsidiary image is henceforth generated independently according to a kind of singular point, respectively. Processing of drawing 8 is common also to S11 of drawing 7, and a terminal point hierarchy picture is also generated through same procedure. Processing by S1 of drawing 6 is completed above.

[0098]In premise art, in order to progress to S2 of drawing 6, matching evaluation is prepared. Drawing 11 shows the procedure. As shown in the figure, two or more valuation plans are set up first (S30). Energy $C^{(m, s)}$, about a pixel introduced by [1.3.2.1] and energy $D^{(m, s)}$, about smoothness of a map introduced by [1.3.2.2] are it. Next, these valuation plans are unified and an overall evaluation type is stood (S31). If eta which total-calories $\lambda C^{(m, s)} + D^{(m, s)}$ introduced by [1.3.2.3] is it, and was introduced by [1.3.2.2] is used, sigmasigma $(\lambda C^{(m, s)}_{(i, j)} + \eta E_0^{(m, s)}_{(i, j)} + E_1^{(m, s)}_{(i, j)})$ (formula 52)

It becomes. However, total calculates by 0, 1 --, and 2^{m-1} about i and j, respectively. Preparation of matching evaluation by the above is completed.

[0099]Drawing 12 is a flow chart which shows details of S2 of drawing 6. As [1] described, matching of a starting point hierarchy picture and a terminal point hierarchy picture is mutually taken by the pictures of the same resolution level. In order to take global matching between pictures good, matching is calculated sequentially from a level with coarse resolution. Since a starting point hierarchy picture and a terminal point hierarchy picture are generated using a singular point filter, a position and luminosity of a singular point are clearly saved also on a level with coarse resolution, and a result of global matching becomes what was dramatically excellent compared with the former.

[0100]Like drawing 12, 0 and the level parameter m are first set as 0 for coefficient-parameters eta (S20). Continue and matching is calculated between each of four subsidiary images of the m-th level in a starting point hierarchy picture, and four subsidiary images of the m-th level in a terminal point hierarchy picture, It asks for 4 kinds of submap $f^{(m, s)}$ ($s= 0, 1, 2, 3$) which fulfills bijection conditions, respectively and makes energy the minimum (S21). Bijection conditions are inspected using an inheritance quadrilateral described by [1.3.3]. Under the present circumstances, since a submap in the m-th level is restrained by them of the m-1st levels as shown in the

formulas 17 and 18, matching in a level with coarser resolution is used one by one. This is the vertical reference between different levels. Although it is $m=0$ now and there is no level coarser than it, this exceptional processing is later mentioned by drawing 13.

[0101] On the other hand, horizontal reference in an identical level is also performed. like the formula 20 of [1.3.3] — $f^{(m, 1)}$ decides $f^{(m, 3)}$ is similar to $f^{(m, 2)}$, and $f^{(m, 2)}$ is similar to $f^{(m, 1)}$, and similar to $f^{(m, 0)}$, respectively. This is because a situation where submaps completely differ is unnatural, if they are contained in the starting point picture same from the first and a terminal point picture at all even if a kind of singular point is different. Energy becomes small, so that submaps are near, as shown in the formula 20, and it considers that matching is good.

[0102] Since there is no submap which can be referred to on level same about $f^{(m, 0)}$ which should be decided first, as shown in the formula 19, a level coarse [one] is referred to. However, in an experiment, after being found to $f^{(m, 3)}$, procedure of updating $f^{(m, 0)}$ once by making this into a constraint was taken. It is equal to this substituting $s=4$ for the formula 20, and making $f^{(m, 4)}$ into new $f^{(m, 0)}$. It is for avoiding a tendency for a degree of association of $f^{(m, 0)}$ and $f^{(m, 3)}$ to become low too much, and an experimental result became fitness more by this measure. In addition to this measure, in an experiment, a shuffle of a submap shown in [1.7.1] was also performed. This is also the meaning which keeps close a degree of association of the submaps originally decided for every kind of singular point. In order to avoid a deviation depending on the starting point of processing, a point of changing a position of the starting point according to a value of s is as [1.7] having described.

[0103] Drawing 13 is a figure showing signs that a submap is determined in the 0th level. Since each subsidiary image merely comprises one pixel, it is automatically decided on the 0th level that all four submap $f^{(0, s)}$ will be unit maps. Drawing 14 is a figure showing signs that a submap is determined in the 1st level. A subsidiary image is constituted from 4 pixels by the 1st level, respectively. These 4 pixels are shown by solid line in the figure. The following procedures are stepped on when looking for corresponding points of the point x of $p^{(1, s)}$ in $q^{(1, s)}$ now.

- [0104] 1. Search for the top left point a of the point x, the upper right point b, the lower left point c, and the lower right point d in resolution of the 1st level.
2. One point a-d looks for a pixel which belongs in a coarse level, i.e., the 0th level. In the case of drawing 14, point a-d belongs to pixel A-D, respectively. However, pixel A-C is a virtual pixel which originally does not exist.
3. Plot corresponding-points A' of pixel A-D which is the 0th level and has already been found — D' in $q^{(1, s)}$. Pixel A' — C' shall be virtual pixels, and shall be in the respectively same position as pixel A-C.
4. Consider that corresponding-points a' of the point a in the pixel A is in pixel A', and plot point a'. At this time, it is assumed that a position which point a' occupies in pixel

A' is the same as a position (in this case, lower right) which the point a occupies in the pixel A.

Corresponding-points b' – d' are plotted by the same method as 5.4, and an inheritance quadrilateral is made from point a' – d'.

6. Look for corresponding-points x' of the point x so that energy may become the minimum in an inheritance quadrilateral. As a candidate of corresponding-points x', the center of a pixel may limit to what is contained in an inheritance quadrilateral, for example. In the case of drawing 14, all of four pixels become a candidate.

[0105]It is a decision procedure of corresponding points of the point x with the above. Same processing is performed about other points of all the, and it opts for a submap. In a level [2nd / more than] level, since it is thought that an inheritance quadrilateral loses its shape gradually, a situation where an interval of pixel A' – D' is vacant as shown in drawing 3 occurs.

[0106]In this way, if four submaps of a certain m-th level are decided, m will be *****ed (S22 of drawing 12), and it will confirm that m is not over n (S23), and will return to S21. It asks for a submap of a level of resolution gradually fine whenever it returns to S21 hereafter, and when it finally returns to S21, map $f^{(n)}$ of the n-th level is decided. Since this map becomes settled about $\eta=0$, it writes it as $f^{(n)}(\eta=0)$.

[0107]Only deltaeta shifts eta in order to ask also for a map about next different eta, and the zero clear of the m is carried out (S24). It checks that new eta is not over predetermined search cancel value η_{max} (S25), returns to S21, and asks for map $f^{(n)}(\eta=\delta\eta)$ about this eta. This processing is repeated and it asks for $f^{(n)}(\eta=i\delta\eta)$ ($i=0, 1, \dots$) by S21. When eta exceeds η_{max} , it progresses to S26, and optimal $\eta=\eta_{opt}$ is determined by a below-mentioned method, and $f^{(n)}(\eta=\eta_{opt})$ is eventually made into map $f^{(n)}$.

[0108]Drawing 15 is a flow chart which shows details of S21 of drawing 12. With this flow chart, a submap in the m-th level is decided about certain fixed eta. When opting for a submap, with premise art, optimal lambda is independently decided for every submap.

[0109]As shown in the figure, the zero clear of s and lambda is carried out first (S210). Next it asks for submap $f^{(m, s)}$ which makes energy the minimum about lambda at that time (it reaches and is about eta to dark) (S211), and this is written to be $f^{(m, s)}(\lambda=0)$. Only deltalambda shifts lambda in order to ask also for a map about different lambda, and it checks that new lambda is not over predetermined search cancel value λ_{max} (S213), returns to S211, and asks for $f^{(m, s)}(\lambda=i\delta\lambda)$ ($i=0, 1, \dots$) by subsequent repetition processings. When lambda exceeds λ_{max} , it progresses to S214, and optimal $\lambda=\lambda_{opt}$ is determined, and $f^{(m, s)}(\lambda=\lambda_{opt})$ is eventually made into map $f^{(m, s)}$ (S214).

[0110]Next, the zero clear of lambda is carried out in order to ask for other submaps

in an identical level, and s is *****ed (S215). It checks that s is not over 4 (S216), and returns to S211. If set to s= 4, like ****, $f^{(m, 0)}$ will be updated using $f^{(m, 3)}$, and determination of a submap in the level will be ended.

[0111] Drawing 16 is a figure showing an action of energy $C^{(m, s)}_f$ corresponding to $f^{(m, s)}$ ($\lambda = \text{idelta}\lambda$) ($i = 0, 1, \dots$) called for while changing λ about certain m and s. If λ increases, $C^{(m, s)}_f$ will usually decrease, as [1.4] described. However, if λ exceeds an optimum value, $C^{(m, s)}_f$ will start to increase. So, with this premise art, λ in case $C^{(m, s)}_f$ takes the minimal value is decided to be λ_{opt} . What is necessary is just to observe the first minimum point, in order for a map to already have collapsed at the time and not to make a meaning, even if $C^{(m, s)}_f$ becomes small again in the range of $\lambda > \lambda_{\text{opt}}$, as shown in the figure. λ_{opt} is independently decided for every submap, and, finally becomes settled one also about $f^{(n)}$.

[0112] On the other hand, drawing 17 is a figure showing an action of energy $C^{(n)}_f$ corresponding to $f^{(n)}$ ($\eta = \text{idelta}\eta$) ($i = 0, 1, \dots$) called for while changing η . If η increases also here, $C^{(n)}_f$ will usually decrease, but if η exceeds an optimum value, $C^{(n)}_f$ will start to increase. Then, η in case $C^{(n)}_f$ takes the minimal value is decided to be η_{opt} . Drawing 17 may consider a figure to which near the zero of a horizontal axis of drawing 4 was expanded. If η_{opt} is decided, terminal decision of the $f^{(n)}$ can be carried out.

[0113] As mentioned above, according to this premise art, various merits are obtained. Since it is not necessary to detect edge probably, a technical problem of edge detection type conventional technology is cancelable. Transcendental knowledge over an object contained in a picture is also unnecessary, and automatic detection of corresponding points realizes it. According to the singular point filter, luminosity and a position of a singular point can be maintained also on a level with coarse resolution, and it is very advantageous to object recognition, feature extraction, and image matching. As a result, a configuration of an image processing system which reduces help work substantially is attained.

[0114] The following modification art is also considered about this premise art.

(1) With premise art, when taking matching between a starting point hierarchy picture and a terminal point hierarchy picture, automatic determination of a parameter was performed, but this method can be generally used, when taking matching between [instead of / between hierarchy pictures] the usual pictures of two sheets.

[0115] For example, between pictures of two sheets, make 2, energy E_0 about a difference of luminosity of a pixel, and energy E_1 about a position gap of a pixel, into a valuation plan, and let these linear combination $E_{\text{tot}} = \alpha E_0 + E_1$ be overall evaluation types. Automatic determination of the α is carried out paying attention to near the extremum of this overall evaluation type. That is, it asks for map that E_{tot} becomes the minimum about various α . α in case E_1 takes the minimal value about α

among those maps is decided to be an optimal parameter. It is considered eventually that a map corresponding to the parameter is the optimal matching between both images.

[0116]What there are various methods in setting out of a valuation plan besides this, for example, takes such a big value like $1/E_1$, and $1/E_2$ that an evaluation result is good may be adopted. An overall evaluation type does not necessarily need to be linear combination, either, and should just choose suitably the n-th power sums ($n=2, 1/2, -1, -2$, etc.), a polynomial, arbitrary functions, etc.

[0117]In the case of [of eta and lambda] 2, any, such as a case beyond it, may be [alpha] sufficient also as a parameter like premise art. When a parameter is three or more, you make it change one [at a time], and it decides.

[0118](2) in this premise art, after deciding a map that a value of an overall evaluation type becomes the minimum, a point that $C^{(m, s)}$, which is one valuation plan which constitutes an overall evaluation type became the minimum was detected, and a parameter was determined. However, it is effective even if it decides a parameter that the minimum of an overall evaluation type only becomes the minimum depending on a situation instead of such two-step time processing. In that case, $\alpha E_0 + \beta E_1$ is made into an overall evaluation type, for example, a constraint which becomes $\alpha + \beta = 1$ may be established and measures, such as treating each valuation plan equally, may be taken. It is because essence of automatic determination of a parameter is in a point of deciding a parameter that energy becomes the minimum.

[0119](3) With premise art, each resolution level generated four kinds of subsidiary images about four kinds of singular points. However, naturally 2 or 3 of 1 and four kinds of kinds may be used selectively. For example, if it is in a state where a bright point exists only one in a picture, a suitable effect should be acquired even if it generates a hierarchy picture only by $f^{(m, 3)}$ about a maximum point. In that case, since a submap which is different with an identical level becomes unnecessary, it is effective in computational complexity about s becoming less.

[0120](4) With this premise art, when one level advanced with a singular point filter, a pixel was set to one fourth. For example, composition which considers it as 1 block by 3x3, and looks for a singular point in it is also possible, and a pixel will be set to one ninth if one level advances in that case.

[0121](5) When a starting point picture and a terminal point picture are colors, change them into a monochrome image first and calculate a map. A color picture of the starting point is changed using a map called for as a result. As the other method, a submap may be calculated about each ingredient of RGB.

[0122][Embodiment about an image effect] An image effect device using the above premise art is explained. This device uses basic technology for recording an animation on a digital camera efficiently, and describes application to a digital camera henceforth [drawing 24] after that explanation. Drawing 18 shows the 1st picture I1 and the 2nd

picture I2 which are key-frames, and a certain pixel $p_1(x_1, y_1)$ and $p_2(x_2, y_2)$ correspond among them. These correspondences are searched for with premise art.

[0123] Drawing 19 cuts a mesh on the 1st picture I1, and shows a correspondence position on the 2nd picture I2 of a polygon which constitutes the mesh. The polygon R1 observed on the 1st picture I1 now was decided by the four lattice points A, B, and C and D. This polygon R1 is named a "starting point polygon." these -- four -- a piece -- a lattice point -- A -- B -- C -- D -- drawing 18 -- having been shown -- as -- respectively -- the -- two -- a picture -- I -- two -- a top -- corresponding -- a point -- A -- ' -- B -- ' -- C -- ' -- D -- ' -- having -- these -- a point -- forming -- having -- a polygon -- R -- two -- " -- a terminal point -- a polygon -- " -- naming . Although a starting point polygon is generally a rectangle, generally a terminal point polygon is a quadrilateral. Anyway, by this embodiment, a pixel unit does not describe correspondence relation between the 1st picture I1 and the 2nd picture I2, but it describes those corresponding picture elements about a lattice point of a starting point polygon. The description is written in a corresponding-points file. By observing a lattice point, capacity of a corresponding-points file can be made very small.

[0124] A corresponding-points file is used in order to generate an intermediate image of the 1st picture I1 and the 2nd picture I2. This point can generate an intermediate image of arbitrary time positions by interpolating a position of corresponding points, as premise art described. Therefore, morphing and a smooth animation between two pictures are generable to it by saving the 1st picture I1, the 2nd picture I2, and a corresponding-points file. For this reason, a compression effect of an animation is acquired.

[0125] Drawing 20 shows how to calculate a correspondence relation about points other than a lattice point from a corresponding-points file. Since there is only information about a lattice point in a corresponding-points file, it is necessary to calculate matching information separately about a point inside a polygon. By drawing 20, correspondence with triangle A'B'C' which hits a lower half of the terminal point polygon R2 similarly is indicated to be triangle ABC which hits a lower half of the starting point polygon R1 of drawing 19. Suppose that the point Q inside triangle ABC by the side of a starting point polygon suited a position which divides the line segment AC internally to t: (1-t), and divides internally between the internally dividing point and point B to s: (1-s) now. What is necessary is just to think that corresponding-points Q' in triangle A'B'C' by the side of a terminal point polygon of this point Q is in a position which divides line segment A'C' internally to t: (1-t), and divides internally between that internally dividing point and point B' to s: (1-s). Namely, what is necessary is to divide a starting point polygon into a triangle and just to search for a point inside a terminal point polygon in the form of internal division of a vector about the triangle. If vector italic shows, it will be $BQ = (1-s) \{ (1-t) BA + t BC \}$.

Come out, and it is, therefore is $B'Q' = (1-s) [(1-t) B'A' + tB'C']$.

Though natural, suppose that the same processing is made also about between the triangle ACD which hits an upper half of the starting point polygon R1, and triangle A'C'D' which hits an upper half of the terminal point polygon R2 similarly.

[0126] Drawing 21 shows the above procedure. A matching result about a lattice point taken on the 1st picture I1 is first acquired like drawing 19 (S10). What is necessary is to perform matching of a pixel unit of premise art and just to extract a portion which hits at a lattice point out of that result at this time. However, a matching result about a lattice point may be specified based on others, and matching art, for example, an optical flow, and block matching, without using premise art.

[0127] It continues and a terminal point polygon is defined on the 2nd picture I2 like the right of drawing 19 (S12). Since a corresponding-points file is above generable, information for identifying the 1st picture I1 and the 2nd picture I2 is incorporated and outputted to this file (S14). Two pictures and a corresponding-points file are held at arbitrary recorders or media, or are put on a direct network or a broadcast wave, and should just be transmitted.

[0128] Drawing 22 shows a procedure which generates an intermediate image using a corresponding-points file. First, the 1st picture I1 and the 2nd picture I2 are read (S20), and a corresponding-points file is read (S22). It continues and correspondence relation between a point in a starting point polygon and a point in a terminal point polygon is computed by a method of drawing 20 (S24). At this time, a correspondence relation about all the pixels within a picture is acquirable. therefore, if coordinates and a color of a point which correspond each other are divided internally by $u: (1-u)$ as premise art described, an intermediate image in a position which divides internally between the 1st picture I1 and the 2nd picture I2 by $:u$ in time $(1-u)$ is generable (S26). Not only interpolation interpolation but extrapolation interpolation may be carried out.

[0129] Drawing 23 shows composition of the image effect device 10 which performs the above processing. The image input part 12 in which this device 10 acquires the 1st picture I1 and the 2nd picture I2 from external memory storage, a photographing camera, etc., The matching processor 14 which performs matching calculation to those pictures by premise art and other art, The corresponding-points file attaching part 16 which stores the corresponding-points file F generated by it, The intermediate image generating part 18 which generates an intermediate image from the 1st picture I1 and the 2nd picture I2, and the indicator 20 displayed as an image near an animation of while carrying out timing adjustment of the 1st picture I1, the 2nd picture I2, and the intermediate image are included. The communications department 22 sends out the 1st picture I1, the 2nd picture I2, and the corresponding-points file F to another network and transmission infrastructure according to a request from the outside. Mesh data which show a size of a mesh, a position of a lattice point, etc. are inputted into the matching processor 14.

[0130]By the above composition, the 1st picture I1 and the 2nd picture I2 which were inputted into the image input part 12 are sent to the matching processor 14. The matching processor 14 performs matching calculation of a pixel unit among those pictures. The matching processor 14 generates the corresponding-points file F based on mesh data, and outputs it to the corresponding-points file attaching part 16.

[0131]The intermediate image generating part 18 responds for a user to ask, or reads the corresponding-points file F by other factors, and generates an intermediate image. This intermediate image is sent to the indicator 20, timing of a generating picture is performed there, and an animation or a morphing picture is displayed. This device 10 may be in the remote terminal side of another composition, that terminal receives comparatively light data called the 1st picture I1, the 2nd picture I2, and the corresponding-points file F in that case, and the intermediate image generating part 18 and the indicator 20 can reproduce an animation independently as this operation also shows.

[0132]The communications department 22 is the composition for it, and it assumes that a remote terminal has already waited at the network point here. That is, the communications department 22 puts the 1st picture I1, the 2nd picture I2, and the corresponding-points file F on a network or a broadcast wave, and sends them out, and a display of an animation is performed by the remote terminal side. Of course, a remote terminal is good for the purpose of not a display but storage. This device inputs the 1st picture I1, the 2nd picture I2, and those corresponding-points files via the communications department 22 from the exteriors, such as a network, passes these to the intermediate image generating part 18, carries out interpolation calculation, and also has a use which generates an intermediate image. The course P of data for it is described by drawing 23.

[0133]When experimented about the contents of processing of an embodiment, very good morphing or an animation compression effect was acquired by adopting about 256x256 size, for example, and setting a lattice point as every direction 10 – every tens of pixel grade as the 1st and 2nd picture. A size of a corresponding-points file is about 10 K bytes from several kilometers, and it has checked that height of image quality and smallness of data volume were realizable.

[0134][Embodiment of a digital camera] Drawing 24 shows composition which applied the image effect device 10 of drawing 23 to the digital camera 50. Hereafter, it explains focusing on a difference with drawing 23. In drawing 24, it changes into the image input part 12, the image pick-up part 52 is formed, and the camera controller 54 which controls this is put side by side. The IC card control section 56 is formed as the corresponding-points file attaching part 16, and this controls input and output of data to IC card 58. Altogether, via the IC card control section 56, the 1st picture I1, the 2nd picture I2, and a corresponding-points file are constituted to an IC card so that writing is possible. Although an output to a network, an external storage, other external

transmission medias, etc. is possible for the communications department 22 in the 1st picture I1, the 2nd picture I2, and a corresponding-points file and it is the composition of receiving data from the IC card control section 56 in the figure, of course, that may not be right and composition received from the usual data bus may be used. The mode setting part 70 sets up photographing mode by the camera controller 54, and specification of "simple animation mode" characteristic of an embodiment besides the usual still picture mode and animation mode is possible for it.

[0135]Drawing 25 shows an example of the image pick-up part 52. An image is acquired by CCD(charge coupled device) 60, is digitized with A/D converter 62, and receives processing of a white balance etc. by the pretreatment part 64 before record. The 1st picture I1 and the 2nd picture I2 of the last form are recorded on IC card 58, or receive processing by the matching processor 14 directly.

[0136]Drawing 26 shows example of another of the image pick-up part 52. Unlike drawing 25, a fixed distance is kept, 2 CCD60 are provided, and photography of a stereo image is possible. A/D converter 62 and the pretreatment part 64 process an image from 2 CCD60 by time sharing. However, also as for these, successive installation **** is also good 2 system.

[0137]Returning to drawing 24, processing of the camera controller 14 is considered as it goes, and it gives an example below.

1. Photo the 1st picture I1 and the 2nd picture I2 with a predetermined time interval (only henceforth a "photographing interval") by the image pick-up part 52 in one photographing operation in the digital camera 50 which adopts ocellus composition of use drawing 25 which compresses an animation as an ocellus camera. At this time, the digital camera 50 is in the above-mentioned simple animation mode, i.e., interim photographing mode of a still picture and an animation.

[0138]If a release button is pushed in this mode for a user's photography, a picture of two sheets will be photoed, for example for a photographing interval in 1 second. When a photographic subject moves in this 1 second or a user moves, a difference comes out between the 1st picture I1 and the 2nd picture I2. The matching processor 14 generates a corresponding-points file in order to bury this difference. An animation corresponding to for 1 second is generable because the intermediate image generating part 18 generates an intermediate image based on this corresponding-points file. If duration time of a reproduction animation is set as more than 1 second by generating many intermediate images, a slow motion picture will also be acquired.

[0139]A generated animation is displayed on the indicator 20 by a liquid crystal device etc., and the user can check the contents of the simple animation. Of course, the indicator 20 may display only the 1st picture I1 and the 2nd picture I2. In any case, since a corresponding-points file is recorded on IC card 58, the animation can express as apparatus of the exterior of the digital camera 50. The external instrument shall have the same composition as the intermediate image generating part 18.

[0140]If a photographing interval in this mode is lengthened, though natural, an animation of longer time is generable. How much time is allowed should just determine by a relation with image quality, and a user may set up. Exposure time may also be decided by the mode setting part 70.

[0141]2. If a photographing interval of the use above which generates a morphing picture as an ocellus camera becomes above to some extent, approach generation of a morphing picture gradually rather than an animation. Therefore, it is good also considering it as specification of the digital camera 50. In this case, further, a view of a photographing interval may be stopped and a user chooses the 1st picture I1 and the 2nd picture I2 with the function of the camera controller 54. Those pictures may be pictures which could choose from an already photoed picture and were inputted from IC card 58. Anyway, morphing is possible also between pictures which are completely unrelated in that case, for example. A morphing picture in which an experiment is also very good is generated.

[0142]3. In the digital camera 50 which adopts compound eye composition of use drawing 26 which generates a multiaspect image as a stereo camera, photo a picture of two sheets simultaneously and the matching processor 14 generates a corresponding-points file. In a corresponding-points file, corresponding points in the 2nd picture I2 photoed by CCD60 of another side are described about each point of the 1st picture I1 photoed by one CCD60. Therefore, depth of the point becomes clear by a principle of triangulation from a horizontal gap of a point (henceforth "a corresponding-points pair") that those pictures correspond each other. As a result, a special effect image is generable by processing of emphasizing depth.

[0143]According to the intermediate image generating part 18, a view image between both CCD60 is generable. If extrapolation interpolation is carried out, a picture from not between CCD60 but a viewpoint distant from the digital camera 50 to some extent is generable. A multiaspect image is also obtained by assuming various viewpoints. This also becomes a basis of a walk-through picture.

[0144]Both CCDboth [one side or] may be provided removable, and it may constitute so that those air clearances can be taken to some extent. In that case, performance as a stereo camera is improved.

[0145]In the above, an example of a digital camera concerning this invention was explained. Although set centering on a personal youth's camera here, this invention may be applied to a camera carried in a professional youth's TV camera and an artificial satellite.

[0146]Like drawing 23, the digital camera 50 inputs the 1st picture I1, the 2nd picture I2, and those corresponding-points files via the communications department 22 or IC card 58 from the exterior, may pass these to the intermediate image generating part 18, may carry out interpolation calculation, and may generate an intermediate image.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] The picture and drawing 1 (c) which are produced by drawing 1 (b) giving an average filter to two persons' person's face, and drawing 1 (d) with drawing 1 (a), the picture, drawing 1 (e), and drawing 1 (f) of $p^{(5,0)}$ which are called for with premise art about two persons' person's face, the picture, drawing 1 (g), and drawing 1 (h) of $p^{(5,1)}$ which are called for with premise art about two persons' person's face, the picture of $p^{(5,2)}$ called for with premise art about two persons' person's face, and drawing 1 (i) and drawing 1 (j) are the photographs of the halftone image which displayed on the display the picture of $p^{(5,3)}$ called for with premise art about two persons' person's face, respectively.

[Drawing 2] The figure, drawing 2 (A), drawing 2 (B), drawing 2 (C), drawing 2 (D), and drawing 2 (E) which drawing 2 (R) shows the quadrilateral of a basis are a figure showing an inheritance quadrilateral, respectively.

[Drawing 3] It is a figure showing the relation between a starting point picture and a terminal point picture, and the relation between the m-th level and the m-1st levels using an inheritance quadrilateral.

[Drawing 4] It is a figure showing the relation between the parameter eta and energy C_f .

[Drawing 5] Drawing 5 (a) and drawing 5 (b) are the figures showing signs that it asks for whether the map about a certain point fulfills bijection conditions from outer product calculation.

[Drawing 6] It is a flow chart which shows the whole premise technical procedure.

[Drawing 7] It is a flow chart which shows the details of S1 of drawing 6.

[Drawing 8] It is a flow chart which shows the details of S10 of drawing 7.

[Drawing 9] It is a figure showing the correspondence relation between a part of picture of the m-th level, and a part of picture of the m-1st levels.

[Drawing 10] It is a figure showing the starting point hierarchy picture generated with premise art.

[Drawing 11] Before progressing to S2 of drawing 6, it is a figure showing the procedure of preparation of matching evaluation.

[Drawing 12] It is a flow chart which shows the details of S2 of drawing 6.

[Drawing 13] It is a figure showing signs that a submap is determined in the 0th level.

[Drawing 14] It is a figure showing signs that a submap is determined in the 1st level.

[Drawing 15] It is a flow chart which shows the details of S21 of drawing 12.

[Drawing 16] It is a figure showing the action of energy $C_f^{(m,s)}$ corresponding to $f^{(m,s)}$ ($\lambda = \text{idelta}\lambda$) called for while changing lambda about certain $f^{(m,s)}$.

[Drawing 17] It is a figure showing the action of energy $C_f^{(n)}$ corresponding to $f^{(n)}$

(eta=ideltaeta) (i= 0, 1, --) called for while changing eta.

[Drawing 18]It is a figure showing signs that the pixel which it is between the 1st picture and the 2nd picture corresponds.

[Drawing 19]It is a figure showing the correspondence relation between the starting point polygon taken on the 1st picture, and the terminal point polygon taken on the 2nd picture.

[Drawing 20]It is a figure showing the procedure of searching for the point in the terminal point polygon corresponding to the point in a starting point polygon.

[Drawing 21]It is a flow chart which shows the procedure which generates a corresponding-points file by an embodiment.

[Drawing 22]It is a flow chart which shows the procedure which generates an intermediate image based on a corresponding-points file by an embodiment.

[Drawing 23]It is a lineblock diagram of the image effect device concerning an embodiment.

[Drawing 24]It is a lineblock diagram of the digital camera concerning an embodiment.

[Drawing 25]It is a lineblock diagram of the image pick-up part of the digital camera shown in drawing 24.

[Drawing 26]It is another lineblock diagram of the image pick-up part of the digital camera shown in drawing 24.

[Description of Notations]

10 Image effect device

12 Image data input part

14 Matching processor

16 Corresponding-points file attaching part

18 Intermediate image generating part

20 Indicator

22 Communications department

50 Digital camera

52 Image pick-up part

54 Camera controller

56 IC card

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-232838

(P2002-232838A)

(43)公開日 平成14年8月16日 (2002.8.16)

(51) Int.Cl. ⁷	識別記号	F I	テマコード*(参考)
H 04 N 5/91		G 03 B 19/02	2 H 05 4
G 03 B 19/02		H 04 N 5/225	F 5 C 02 2
H 04 N 5/225		5/232	Z 5 C 02 3
5/232		5/262	5 C 05 3
5/262		13/02	5 C 06 1

審査請求 有 請求項の数14 O L (全 25 頁) 最終頁に続く

(21)出願番号 特願2001-110393(P2001-110393)
(22)出願日 平成13年4月9日(2001.4.9)
(31)優先権主張番号 特願2000-360638(P2000-360638)
(32)優先日 平成12年11月28日(2000.11.28)
(33)優先権主張国 日本 (JP)

(71)出願人 396001980
株式会社モノリス
東京都港区麻布十番1丁目7番3号
(72)発明者 秋吉 仰三
東京都港区麻布十番1-7-3 株式会社
モノリス内
(72)発明者 秋吉 信雄
東京都港区麻布十番1-7-3 株式会社
モノリス内
(72)発明者 品川 嘉久
東京都江戸川区西葛西5-10-26
(74)代理人 100105924
弁理士 森下 賢樹

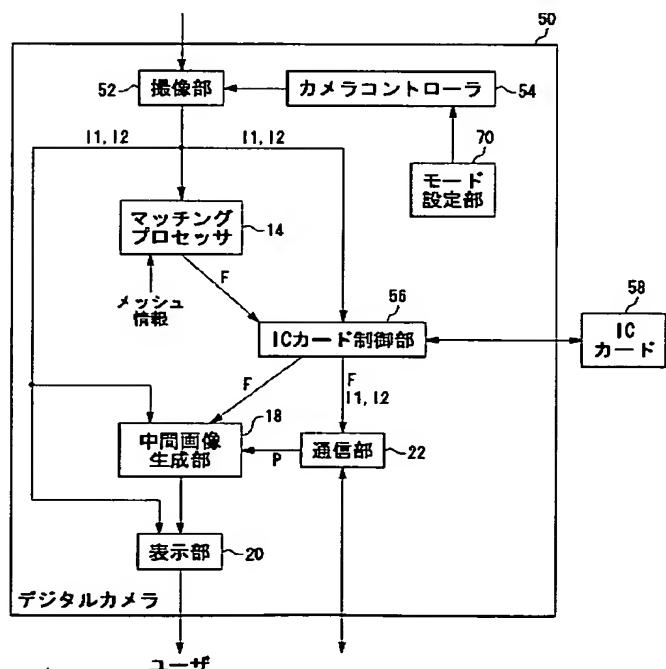
最終頁に続く

(54)【発明の名称】デジタルカメラ

(57)【要約】

【課題】 デジタルカメラはメモリ容量に限界があり、動画の圧縮率が問題になる。

【解決手段】 撮像部52は第1画像と第2画像を入力する。マッチングプロセッサ14はそれらの画像間で画素マッチングを行い、第1画像上に取られたメッシュの格子点について、その対応点を第2画像上に求める。その結果が対応点ファイルFとして記録される。中間画像生成部18は、対応点ファイルFをもとに第1画像と第2画像間の中間画像を生成する。これが動画を近似する。格子点だけで対応点ファイルを作ったので、データ容量は非常に小さくなる。



【特許請求の範囲】

【請求項1】 画像を撮影する撮像部と、
所定の間隔で撮像部に第1および第2画像を撮影させる
カメラコントローラと、
第1画像および第2画像の間でマッチングを計算し、そ
の結果を対応点ファイルとして出力するマッチングプロ
セッサと、
を含むことを特徴とするデジタルカメラ。

【請求項2】 画像を撮影する撮像部と、
撮像部によって撮影された画像のうち、2枚の画像を第
1および第2画像として決定するカメラコントローラ
と、
第1画像および第2画像の間でマッチングを計算し、そ
の結果を対応点ファイルとして出力するマッチングプロ
セッサと、
を含むことを特徴とするデジタルカメラ。

【請求項3】 ステレオ視を実現する撮像部と、
ステレオ画像を構成する第1および第2画像を撮像部に
撮影させるカメラコントローラと、
第1画像および第2画像の間でマッチングを計算し、そ
の結果を対応点ファイルとして出力するマッチングプロ
セッサと、
を含むことを特徴とするデジタルカメラ。

【請求項4】 前記対応点ファイルをもとに、第1画像
および第2画像の中間画像を生成する中間画像生成部を
さらに含む請求項1から3のいずれかに記載のデジタル
カメラ。

【請求項5】 第1画像、第2画像および中間画像を表
示する表示部をさらに含む請求項4に記載のデジタルカ
メラ。

【請求項6】 第1画像、第2画像および対応点ファ
イルを関連づけて記録する対応点ファイル保持部をさら
に含む請求項4、5のいずれかに記載のデジタルカメラ。

【請求項7】 前記マッチングプロセッサは、第1画像
上に設けられたメッシュの格子点に対応する第2画像上
の点を画像マッチングによって検出し、その結果をもと
に、第1画像上のメッシュを構成する起点ポリゴンに、
第2画像上にて対応する終点ポリゴンを定義する請求項
1から6のいずれかに記載のデジタルカメラ。

【請求項8】 前記マッチングプロセッサは、第1画像
について二次元的な探索を行って検出した特異点と、第
2画像について二次元的な探索を行って検出した特異点
との対応をもとに画素単位のマッチング計算を行う請求
項7に記載のデジタルカメラ。

【請求項9】 前記マッチングプロセッサは、第1画像
と第2画像を、それぞれ前記特異点を抽出することによ
って多重解像度化したうえで、同一解像度レベル間で
画像単位のマッチング計算を行い、その結果を異なる解
像度レベルにおけるマッチング計算に継承しながら、最終
的に最も解像度の細かいレベルにおける画素単位の対応

関係を取得する請求項8に記載のデジタルカメラ。

【請求項10】 前記撮像部による簡易動画撮影のモー
ドを設定するためのモード設定部をさらに有する請求項
1に記載のデジタルカメラ。

【請求項11】 第1画像および第2画像を取得する撮
像部と、

第1画像と第2画像のマッチングを計算するマッチング
プロセッサとを含み、

このマッチングプロセッサは、第1画像上のメッシュを
構成する起点ポリゴンに、第2画像上にて対応する終点
ポリゴンを定義することにより、対応点ファイルを生成
することを特徴とするデジタルカメラ。

【請求項12】 画像を撮影する撮像部と、
所定の間隔で撮像部に第1および第2画像を撮影させる
カメラコントローラと、

第1画像および第2画像の間でマッチングを計算し、そ
の結果を対応点ファイルとして出力するマッチングプロ
セッサとを含み、

このマッチングプロセッサは、第1画像と第2画像のそ
れぞれを特異点を抽出することによって階層化し、その
階層の粗い方から順次特異点どうしの対応関係を検出
することにより、最終的に最も細かい階層に位置する第1
画像と第2画像の対応を検出することを特徴とするデジ
タルカメラ。

【請求項13】 画像を撮影する撮像部と、
撮像部によって撮影された画像のうち、2枚の画像を第
1および第2画像として決定するカメラコントローラ
と、

第1画像および第2画像の間でマッチングを計算し、そ
の結果を対応点ファイルとして出力するマッチングプロ
セッサとを含み、

このマッチングプロセッサは、第1画像と第2画像のそ
れぞれを特異点を抽出することによって階層化し、その
階層の粗い方から順次特異点どうしの対応関係を検出
することにより、最終的に最も細かい階層に位置する第1
画像と第2画像の対応を検出することを特徴とするデジ
タルカメラ。

【請求項14】 ステレオ視を実現する撮像部と、
ステレオ画像を構成する第1および第2画像を撮像部に
撮影させるカメラコントローラと、

第1画像および第2画像の間でマッチングを計算し、そ
の結果を対応点ファイルとして出力するマッチングプロ
セッサとを含み、

このマッチングプロセッサは、第1画像と第2画像のそ
れぞれを特異点を抽出することによって階層化し、その
階層の粗い方から順次特異点どうしの対応関係を検出
することにより、最終的に最も細かい階層に位置する第1
画像と第2画像の対応を検出することを特徴とするデジ
タルカメラ。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、デジタルカメラに関し、とくに撮影した画像に所定の処理をなすデジタルカメラに関する。

【0002】

【従来の技術】B S デジタル放送が開始され、また多数のユーザが P C や形態電話からインターネット上のサービスを享受するようになり、放送と通信の垣根が急速に崩れつつある。さらには、ビデオ機器やカメラがデジタル化し、個人ユースの情報機器が放送や通信と非常に高い親和性を有するようになった。一時は概念が見えにくいといわれた「マルチメディア」は、21世紀を迎えた今日、技術革新とインフラの整備により、人類文化のひとつの中核となった。

【0003】当初は画像のデジタル化による効率的な保存と印刷を目して登場したデジタルカメラも、今日では各種画像処理技術を搭載し、パーソナル色の強化とともに、プロフェッショナルユースにも耐える機能を実現しつつある。個人主導のデジタル機器が今後 I T 文化的進展を加速することに疑いを差し挟む余地はない。

【0004】

【発明が解決しようとする課題】一部のデジタルカメラには、ハイパスフィルタによるエッジ強調や、色調変換処理など、ユーザの嗜好に合った画像を生成する技術が利用されている。また、動画をなるべく長時間撮影すべく、M P E G (Motion Picture Expert Group) による圧縮などの技術が採用されている。こうした状況にあって、本発明の目的は、新たな画像処理の採用により、動画を比較的軽いデータで実現するデジタルカメラの提供にある。

【0005】

【課題を解決するための手段】本発明のある態様はデジタルカメラに関する。このカメラは、画像マッチングを時間について利用する。すなわち、このカメラは画像を撮影する撮像部と、所定の間隔で撮像部に第1および第2画像を撮影させるカメラコントローラと、第1画像および第2画像の間でマッチングを計算し、その結果を対応点ファイルとして出力するマッチングプロセッサなどを含む。「所定の間隔」は、ユーザが設定可能でもよく、予め固定されていてもよい。

【0006】カメラコントローラは、例えばユーザが一枚の画像の撮影を指示したとき、この所定の間隔で2枚の画像を順に撮影する。マッチングプロセッサは、2枚の画像のマッチングをもとに対応点ファイルを作るために、後にこのファイルを用いて2枚の画像の中間画像を生成できる。その結果、少ないデータ量で簡易的に動画を再現することができる。2枚の画像の撮影間隔がある程度開ければ、動画の再生よりも、むしろ映像効果でいうモーフィング画像が得られる。それもカメラの機能として非常に面白い。例えば、2枚の画像を別の2人の顔

とすれば、それらの顔のモーフィング画像を生成することができます。

【0007】本発明の別の態様もデジタルカメラに関する。このカメラは、画像を撮影する撮像部と、撮像部によって撮影された画像のうち、二枚の画像を第1および第2画像として決定するカメラコントローラと、第1画像および第2画像の間でマッチングを計算し、その結果を対応点ファイルとして出力するマッチングプロセッサなどを含む。カメラコントローラは、ユーザの指示にしたがい、または所定の規則に従い、撮像部が撮影した画像のうち、所定の2枚の画像を第1および第2画像として決定してもよい。この態様によれば、より高い自由度で、前述のモーフィング画像、または圧縮動画を得ることができる。この態様は、画像マッチングを時間もしくは空間、または両者について利用する。

【0008】本発明の別の態様もデジタルカメラに関する。このカメラは、画像マッチングを空間について利用する。すなわち、このカメラはステレオ視を実現する撮像部と、撮像部によってステレオ画像を構成する第1および第2画像を撮影させるカメラコントローラと、第1画像および第2画像の間でマッチングを計算し、その結果を対応点ファイルとして出力するマッチングプロセッサなどを含む。したがって、この対応点ファイルをもとに、特殊効果画像や視点変更画像を生成することができる。ステレオ画像の対応点をもとに、その画像の各点の奥行き情報が判明するためである。

【0009】このカメラはさらに、対応点ファイルをもとに、第1画像および第2画像の中間画像を生成する中間画像生成部を含んでもよい。中間画像は時間または空間、場合により、その両者に関する補間画像である。第1画像、第2画像および中間画像を動画や中間視点画像などとして表示する表示部をさらに備えてよい。また、第1画像、第2画像および対応点ファイルを関連づけて記録する対応点ファイル保持部、例えば I C カードその他のメモリカード、またはそれらの制御回路をさらに含んでもよい。

【0010】マッチングプロセッサは、第1画像上に設けられたメッシュの格子点に対応する第2画像上の点を画像マッチングによって検出し、その結果をもとに、第1画像上のメッシュを構成する起点ポリゴンに、第2画像上にて対応する終点ポリゴンを定義してもよい。マッチングプロセッサは、第1画像と第2画像の間で画素単位のマッチング計算を行ってよい。この計算は、全画素について行ってよいし、格子点のみ、または格子点とそれに関連する画素について行ってよい。

【0011】またマッチングプロセッサは、第1画像について二次元的な探索を行って検出した特異点と、第2画像について二次元的な探索を行って検出した特異点との対応をもとに画素単位のマッチング計算を行ってよい。この場合さらに、第1画像と第2画像を、それぞれ

特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得してもよい。

【0012】前述の特異点を用いるマッチング方法は、本出願人が先に特許第2927350号にて提案した技術（以下「前提技術」という）の応用であり、マッチングプロセッサの処理に好適である。ただし、前提技術は格子点またはそれによって決まるポリゴンに関する本発明の特徴にはいっさい触れていない。本発明でポリゴンという一種の簡略化技術を導入したため、前記対応点ファイルのサイズを非常に小さくすることができる。

【0013】すなわち、第1、第2画像がそれぞれ $n \times m$ 画素を有する場合、それらの画素単位の対応をそのまま記述すると、 $(n \times m)^2$ という組合せが生じ、対応点ファイルが膨大なデータ量になる。しかし、これを格子点どうしの対応関係、または実質的に同じことであるが、格子点で決まるポリゴンどうしの対応関係を記述することに改めることでデータ量が大幅に減る。第1、第2画像および対応点ファイルのみを保有すれば、動画を再現でき、動画の伝送、ストレージ等において大きな効果を發揮する。画像の記録容量に制限のあるデジタルカメラに好適な技術である。

【0014】なお、本発明は前提技術を必須とはしない。また、以上の各構成、工程を任意に入れ替えたり、方法と装置の間で表現を一部または全部入れ替え、または追加したり、表現をコンピュータプログラム、記録媒体等に変更したものもまた、本発明として有効である。

【0015】

【発明の実施の形態】はじめに、実施の形態で利用する多重解像度特異点フィルタ技術とそれを用いた画像マッチング処理を「前提技術」として詳述する。これらの技術は本出願人がすでに特許第2927350号を得ている技術であり、本発明との組合せに最適である。本発明では、画像上にメッシュを設け、その格子点によって多数の画素を代表されるため、もともと前提技術のような画素単位のマッチング技術に対する適用効果が高いためである。ただし、実施の形態で採用可能な画像マッチング技術はこれに限られない。図18以降、前提技術を利用したデジタルカメラを具体的に説明する。

$$\begin{aligned} p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\ p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\ p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\ p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)})) \end{aligned}$$

【0016】[前提技術の背景]最初に [1] で前提技術の要素技術の詳述し、[2] で処理手順を具体的に説明する。さらに [3] で実験の結果を報告する。

【0017】[1] 要素技術の詳細

[1. 1] イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから精細なレベルへと順に解像度の階層を辿っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

【0018】本前提技術は、例えば完全に自動的なモーフィング、物体認識、立体写真測量、ボリュームレンダリング、少ないフレームからの滑らかな動画像の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ボリュームレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

【0019】[1. 2] 特異点フィルタの階層

前提技術に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の輝度及び位置を保存することができる。ここで画像の幅を N 、高さを M とする。以下簡単のため、 $N = M = 2^n$ (n は自然数) と仮定する。また、区間 $[0, N] \subset \mathbb{R}$ を I と記述する。 (i, j) における画像の画素を $p_{(i,j)}$ と記述する ($i, j \in I$)。

【0020】ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第 m レベルにおける各画像のサイズは $2^m \times 2^m$ ($0 \leq m \leq n$) とする。特異点フィルタは次の4種類の新たな階層画像を n から下がる方向で再帰的に構築する。

【0021】

【数1】

ただしここで、

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)}$$

とする。以降これら4つの画像を副画像（サブイメージ）と呼ぶ。 $m \leq n \leq t \leq x+1$ 、 $\max_{x \leq t \leq x+1} \alpha$ 及び β をそれぞれ記述すると、副画像はそれ以下のように記述できる。

【0022】

$$p_{(m,0)} = \alpha(x) \alpha(y) p_{(m+1,0)}$$

$$p_{(m,1)} = \alpha(x) \beta(y) p_{(m+1,1)}$$

$$p_{(m,2)} = \beta(x) \alpha(y) p_{(m+1,2)}$$

$$p_{(m,3)} = \beta(x) \beta(y) p_{(m+1,3)}$$

すなわち、これらは α と β のテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明らかなように、特異点フィルタはもとの画像について 2×2 画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、前提技術では輝度を採用するが、画像に関するいろいろな数値を採用することができる。ふたつの方向の両方について最大画素値となる画素は極大点、ふたつの方向の両方について最小画素値となる画素は極小点、ふたつの方向の一方について最大画素値となるとともに、他方について最小画素値となる画素は鞍点として検出される。

【0023】特異点フィルタは、各ブロックの内部で検出された特異点の画像（ここでは1画素）でそのブロックの画像（ここでは4画素）を代表させることにより、画像の解像度を落とす。特異点の理論的な観点からすれば、 $\alpha(x) \alpha(y)$ は極小点を保存し、 $\beta(x) \beta(y)$ は極大点を保存し、 $\alpha(x) \beta(y)$ 及び $\beta(x) \alpha(y)$ は鞍点を保存する。

【0024】はじめに、マッチングをとるべき始点（ソース）画像と終点（デスティネーション）画像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

【0025】この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとれしていく。まず $p_{(m,0)}$ を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p_{(m,1)}$ を用いて鞍点のマッチングがとられ、 $p_{(m,2)}$ を用いて他の鞍

$$f^{(m,i)} : p^{(m,\sigma(i))} \rightarrow q^{(m,\sigma(i))}$$

ここで $\sigma(i) \in \{0, 1, 2, 3\}$ である。

【003.0】 [1. 3. 1] 全単射

始点画像と終点画像の間のマッチングを写像で表現する場合、その写像は両画像間で全単射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射

【数2】

$$= p_{(i,j)} \quad (\text{式 } 2)$$

点のマッチングがとられる。そして最後に $p_{(m,3)}$ を用いて極大点のマッチングがとられる。

【0026】図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像 $p_{(5,0)}$ を示している。同様に、図1(e)と図1(f)は $p_{(5,1)}$ 、図1(g)と図1(h)は $p_{(5,2)}$ 、図1(i)と図1(j)は $p_{(5,3)}$ をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず $p_{(5,0)}$ によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p_{(5,1)}$ によれば口が明確になる。口は横方向で輝度が低いためである。 $p_{(5,2)}$ によれば首の両側の縦線が明確になる。最後に、 $p_{(5,3)}$ によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

【0027】特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影された画像の特徴と、予め記録しておいたいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

【0028】 [1. 3] 画像間の写像の計算

始点画像の位置 (i, j) の画素を $p_{(n,i,j)}$ と書き、同じく終点画像の位置 (k, l) の画素を $q_{(n,k,l)}$ で記述する。 $i, j, k, l \in I$ とする。画像間の写像のエネルギー（後述）を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ $p_{(m,0)}$ と $q_{(m,0)}$ 間の写像 $f_{(m,0)} : p_{(m,0)} \rightarrow q_{(m,0)}$ が計算される。 $f_{(m,0)}$ に基づき、最小エネルギーを持つ $p_{(m,1)}$ 、 $q_{(m,1)}$ 間の写像 $f_{(m,1)}$ が計算される。この手続は、 $p_{(m,3)}$ と $q_{(m,3)}$ の間の写像 $f_{(m,3)}$ の計算が終了するまで続く。各写像 $f_{(m,i)} (i = 0, 1, 2, \dots)$ を副写像と呼ぶことにする。 $f_{(m,i)}$ の計算の都合のために、 i の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

【0029】

【数3】

$$(\text{式 } 3)$$

かつ单射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全单射のディジタル版である。前提技術では、画素は格子点によって特定される。

【0031】始点副画像（始点画像について設けられた

副画像) から終点副画像(終点画像について設けられた副画像)への写像は、 $f^{(m, s)} : I / 2^n - m \times I / 2^n - m \rightarrow I / 2^n - m \times I / 2^n - m$ ($s = 0, 1, \dots$) によって表される。ここで、 $f^{(m, s)}(i, j) = (k, l)$ は、始点画像の $p^{(m, s)}(i, j)$ が終点画像の $q^{(m, s)}(k, l)$ に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$ が成り立つとき画素 $q^{(k, l)}$ を $q_f(i, j)$ と記述する。

【0032】前提技術で扱う画素(格子点)のようにデータが離散的な場合、全単射の定義は重要である。ここでは以下のように定義する (i, i', j, j', k ,

$$\overrightarrow{p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)}} \quad \overrightarrow{p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)}} \quad \overrightarrow{p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{p_{(i,j+1)}^{(m,s)} p_{(i,j)}^{(m,s)}} \quad (式5)$$

この正方形は写像 f によって終点画像平面における四辺形に写像されなければならない。 $f^{(m, s)}(R)$ に

$$q_f^{(m,s)}(i,j) q_f^{(m,s)}(i+1,j) q_f^{(m,s)}(i+1,j+1) q_f^{(m,s)}(i+1,j+1)$$

は、以下の全単射条件を満たす必要がある。

- 【0034】1. 四辺形 $f^{(m, s)}(R)$ のエッジは互いに交差しない。
- 2. $f^{(m, s)}(R)$ のエッジの方向は R のそれらに等しい(図2の場合、時計回り)。
- 3. 緩和条件として収縮写像(リトラクション: retractions)を許す。

【0035】何らかの緩和条件を設けないかぎり、全単射条件を完全に満たす写像は単位写像しかないとある。ここでは $f^{(m, s)}(R)$ のひとつのエッジの長さが0、すなわち $f^{(m, s)}(R)$ は三角形になってもよい。しかし、面積が0となるような图形、すなわち1点または1本の線分になってはならない。図2(R)がもとの四辺形の場合、図2(A)と図2(D)は全単射条件を満たすが、図2(B)、図2(C)、図2(E)は満たさない。

【0036】実際のインプリメンテーションでは、写像

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_f^{(m,s)}(i,j))|^2$$

ここで、 $V(p^{(m, s)}(i, j))$ 及び $V(q^{(m, s)}_{f(i, j)})$ はそれぞれ画素 $p^{(m, s)}(i, j)$ 及び $q^{(m, s)}_{f(i, j)}$ の輝度である。 f のトータルのエネルギー $C^{(m, s)}$ は、マッチ

$$C_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} C_{(i,j)}^{(m,s)}$$

【1. 3. 2. 2】滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギー D_f を導入する。このエネルギーは画素の輝度とは関

係なく、 $p^{(m, s)}(i, j)$ および $q^{(m, s)}_{f(i, j)}$ の位置によって決まる ($i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$)。点 (i, j) における写像 $f^{(m, s)}$ のエネルギー $D^{(m, s)}(i, j)$

【数4】

$$p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)} \quad (式4)$$

を考える ($i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$)。ここで R の各辺(エッジ)の方向を以下のように定める。

【0033】

【数5】

よって示される四辺形、

【数6】

(式6)

が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというのである。すなわち、 $f(i, j) = (i, j)$ (ただし $i = 0, i = 2^m - 1, j = 0, j = 2^m - 1$ の4本の線上) である。この条件を以下「附加条件」とも呼ぶ。

【0037】[1. 3. 2] 写像のエネルギー

[1. 3. 2. 1] 画素の輝度に関するコスト
写像 f のエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像 $f^{(m, s)}$ の点 (i, j) におけるエネルギー $C^{(m, s)}(i, j)$ は次式によって定まる。

【0038】

【数7】

(式7)

ングを評価するひとつの評価式であり、つぎに示す $C^{(m, s)}(i, j)$ の合計で定義できる。

【0039】

【数8】

(式8)

係なく、 $p^{(m, s)}(i, j)$ および $q^{(m, s)}_{f(i, j)}$ の位置によって決まる ($i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$)。点 (i, j) における写像 $f^{(m, s)}$ のエネルギー $D^{(m, s)}(i, j)$

は次式で定義される。

【0040】

$$D_{(i,j)}^{(m,s)} = \eta E_0^{(m,s)} + E_1^{(m,s)}$$

ただし、係数パラメータ η は 0 以上の実数であり、また、

$$E_0^{(m,s)} = \|(i,j) - f^{(m,s)}(i,j)\|^2 \quad (\text{式 } 10)$$

【数11】

$$E_1^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2 / 4 \quad (\text{式 } 11)$$

とする。ここで、

$$\|(x,y)\| = \sqrt{x^2 + y^2}$$

であり、 $i' < 0$ および $j' < 0$ に対して $f(i', j')$ は 0 と決める。 E_0 は (i, j) 及び $f(i, j)$ の距離で決まる。 E_1 は画素があまりにも離れた画素へ写影されることを防ぐ。ただし E_0 は、後に別のエネルギー関数で置き換える。 E_1 は写像の滑らかさを保証する。 E_1 は、 $p_{(i,j)}$ の変位とその隣接点の変

$$D_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} D_{(i,j)}^{(m,s)} \quad (\text{式 } 13)$$

[1. 3. 2. 3] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は $\lambda C_f^{(m,s)} + D_f^{(m,s)}$ で定義される。ここで係数パラメータ λ は 0 以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次式で示す最小エネルギーを与える写像を見いだすことである。

【0042】

【数14】

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \quad (\text{式 } 14)$$

$\lambda = 0$ 及び $\eta = 0$ の場合、写像は単位写像になることに注意すべきである（すなわち、全ての $i = 0, \dots, 2^m - 1$ 及び $j = 0, \dots, 2^m - 1$ に対して $f^{(m,s)}(i, j) = (i, j)$ となる）。後述のごとく、本前提技術では最初に $\lambda = 0$ 及び $\eta = 0$ の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の λ の位置を変えて $C_f^{(m,s)} + \lambda D_f^{(m,s)}$ と定義したとすれば、 $\lambda = 0$ 及び $\eta = 0$ の場合に総合評価式が $C_f^{(m,s)}$ だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけで対応づけられ、写像が無意味なものになる。こうした無意味な写像をもとに写像を変形して

【数9】

【数10】

（式9）

【数12】

（式12）

位の間の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー D_f は次式で定まる。

【0041】

【数13】

いってもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

【0043】オプティカルフローもこの前提技術同様、画素の輝度の差と滑らかさを考慮する。しかし、オプティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。前提技術に係る特異点フィルタを用いることによって大域的な対応関係を検出することができる。

【0044】 [1. 3. 3] 多重解像度の導入による写像の決定

最小エネルギーを与え、全単射条件を満足する写像 $f_{m+1,n}$ を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0045】まず、

【数15】

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}])$$

が成り立つとき、 $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$ をそれぞれ $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$ のparentと呼ぶことにする。[x]はxを越えない最大整数である。また $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$ をそれぞれ

$$\text{parent}(i, j) = ([\frac{i}{2}], [\frac{j}{2}])$$

$p^{(m, s)}(i, j)$ と $q^{(m, s)}(k, l)$ の間の写像 $f^{(m, s)}$ は、エネルギー計算を行って最小になったものを見つけることで決定される。 $f^{(m, s)}(i, j) = (k, l)$ の値は $f^{(m-1, s)}(m=1, 2, \dots, n)$ を用いることによって、以下のように

$$q_{g^{(m,s)}}^{(m,s)}(i-1, j-1), q_{g^{(m,s)}}^{(m,s)}(i-1, j+1), q_{g^{(m,s)}}^{(m,s)}(i+1, j+1), q_{g^{(m,s)}}^{(m,s)}(i+1, j-1)$$

ただしここで、

$$g^{(m,s)}(i, j) = f^{(m-1,s)}(\text{parent}(i, j)) + f^{(m-1,s)}(\text{parent}(i, j) + (1, 1))$$

である。こうして定めた四辺形を、以下 $p^{(m, s)}(i, j)$ の相続(inherited)四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0048】図3は以上の手順を示している。同図において、始点画像のA, B, C, Dの画素は、第m-1レベルにおいてそれぞれ終点画像のA', B', C', D'へ写影される。画素 $p^{(m, s)}(i, j)$ は、相続四辺形A' B' C' D'の内部に存在する画素q

$$E_{0(i,j)} = \|f^{(m,0)}(i, j) - g^{(m)}(i, j)\|^2$$

また、副写像 $f^{(m, s)}$ を計算するためには次式を用いる。

$$E_{0(i,j)} = \|f^{(m,s)}(i, j) - f^{(m,s-1)}(i, j)\|^2 \quad (1 \leq i)$$

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式20により、異なる特異点に対応する副写像が、副写像どうしの類似度が高くなるように同一レベル内で関連づけられる。式19は、 $f^{(m, s)}(i, j)$ と、第m-1レベルの画素の一部と考えた場合の(i, j)が射影されるべき点の位置との距離を示している。

【0052】仮に、相続四辺形A' B' C' D'の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、A' B' C' D'の境界線からの距離がL (始めはL=1)である画素を調べる。それらのう

(式15)

$p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$ のchildと呼ぶ。関数parent(i, j)は次式で定義される。

【0046】

【数16】

(式16)

決定される。まず、 $q^{(m, s)}(k, l)$ は次の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

【0047】

【数17】

(式17)

【数18】

(式18)

$f^{(m, s)}(i, j)$ へ写影されなければならない。以上の配慮により、第m-1レベルの写像から第mレベルの写像への橋渡しがなされる。

【0049】先に定義したエネルギー E_0 は、第mレベルにおける副写像 $f^{(m, 0)}$ を計算するために、次式に置き換える。

【0050】

【数19】

(式19)

【0051】

【数20】

(式20)

ち、エネルギーが最小になるものが全単射条件を満たせば、これを $f^{(m, s)}(i, j)$ の値として選択する。そのような点が発見されるか、またはしがその上限の $L^{(m)} \max$ に到達するまで、Lを大きくしていく。 $L^{(m)} \max$ は各レベルmに対して固定である。そのような点が全く発見されない場合、全単射の第3の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f^{(m, s)}(i, j)$ を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の第1及び第2条件を外す。

【0053】多重解像度を用いる近似法は、写像が画像

の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことができない。さらに、通常写像に滑らかさを要求するため、そうした画素間の対応関係を見つけにくくしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、そうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル（粗いレベル）において小さいためである。

【0054】[1. 4] 最適なパラメータ値の自動決定
既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。前提技術に係る方法によれば、最適なパラメータ値を完全に自動決定することができる。

【0055】 前提技術に係るシステムはふたつのパラメータ、 λ 及び η を含む。端的にいえば、 λ は画素の輝度の差の重みであり、 η は写像の剛性を示している。これらのパラメータの値は初期値が0であり、まず $\eta=0$ に固定して λ を0から徐々に増加させる。 λ の値を大きくしながら、しかも総合評価式（式14）の値を最小にする場合、各副写像に関する $C_{(m, s)}^f$ の値は一般に小さくなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、 λ が最適値を超えると以下の現象が発生する。

【0056】1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。

2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる。

【0057】3. その結果、式14において $D_{(m, s)}^f$ が急激に増加しようとする。

4. その結果、式14の値が急激に増加しようとするた

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{l=\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

で示されるA個の画素が、

$$C_f^{(m, s)} - l^2 = C_f^{(m, s)} - \frac{1}{\lambda}$$

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素のエネルギーがすべてゼロにな

$$\partial C_f^{(m, s)} = -\frac{A}{\lambda}$$

だけ変化することを示し、その結果、

め、 $D_{(m, s)}^f$ の急激な増加を抑制するよう $f_{(m, s)}$ が変化し、その結果 $C_{(m, s)}^f$ が増加する。

したがって、 λ を増加させながら式14が最小値をとるという状態を維持しつつ $C_{(m, s)}^f$ が減少から増加に転じる閾値を検出し、その λ を $\eta=0$ における最適値とする。つぎに η を少しづつ増やして $C_{(m, s)}^f$ の挙動を検査し、後述の方法で η を自動決定する。その η に対応して λ も決まる。

【0058】 この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できるとき、その目が固定される。

【0059】[1. 4. 1] λ の動的決定

λ は0から所定の刻み幅で増加されていく、 λ の値が変わることで副写像が評価される。式14のごとく、総エネルギーは $\lambda C_{(m, s)}^f + D_{(m, s)}^f$ によって定義される。式9の $D_{(m, s)}^f$ は滑らかさを表すもので、理論的には単位写像の場合に最小になり、写像が歪むほど E_0 も E_1 も増加していく。 E_1 は整数であるから、 $D_{(m, s)}^f$ の最小刻み幅は1である。このため、現在の $\lambda C_{(m, s)}^f(i, j)$ の変化（減少量）が1以上でなければ、写像を変化させることによって総エネルギーを減らすことはできない。なぜなら、写像の変化に伴って $D_{(m, s)}^f$ は1以上増加するため、 $\lambda C_{(m, s)}^f(i, j)$ が1以上減少しない限り総エネルギーは減らないためである。

【0060】 この条件のもと、 λ の増加に伴い、正常な場合に $C_{(m, s)}^f(i, j)$ が減少することを示す。 $C_{(m, s)}^f(i, j)$ のヒストグラムを $h(1)$ と記述する。 $h(1)$ はエネルギー $C_{(m, s)}^f(i, j)$ が 1^2 である画素の数である。 $\lambda 1^2 \geq 1$ が成立るために、例えば $1^2 = 1/\lambda$ の場合を考える。 λ が λ_1 から λ_2 まで微小量変化するとき、

【数21】

(式21)

【数22】

(式22)

ると近似している。この式は $C_{(m, s)}^f$ の値が、

【数23】

(式23)

【数24】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (\text{式 2 4})$$

が成立する。 $h(l) > 0$ であるから、通常 $C_{(m,s)}$ は減少する。しかし、 λ が最適値を越えようとするとき、上述の現象、つまり $C_{(m,s)}$ の増加が発生する。この現象を検出することにより、 λ の最適値を決定する。

【0061】なお、 H ($h > 0$) 及び k を定数とすると、

$$C_f^{(m,s)} = C + \frac{H}{(3/2 + k/2)\lambda^{3/2+k/2}} \quad (\text{式 2 7})$$

となる。これが $C_{(m,s)}$ の一般式である (C は定数)。

【0062】 λ の最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率を p_0 と仮定する。この場合、

【数 28】

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式 2 8})$$

が成立しているため、全単射条件を破る画素の数は次式の率で増加する。

【0063】

【数 29】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}} \quad (\text{式 2 9})$$

従って、

【数 30】

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1 \quad (\text{式 3 0})$$

は定数である。仮に $h(l) = H l^k$ を仮定するとき、例えば、

【数 31】

$B_0 \lambda^{3/2+k/2} = p_0 H$ (式 3 1)
は定数になる。しかし λ が最適値を越えると、上の値は急速に増加する。この現象を検出し、 $B_0 \lambda^{3/2+k/2}$ の値が異常値 $B_{\text{threshold}}$ を越えるかどうかを検査し、 λ の最適値を決定することができる。同様に、 $B_1 \lambda^{3/2+k/2}$ の値が異常値 $B_{1\text{ threshold}}$ を越えるかどうかを検査することにより、全単射の第3の条件を破る画素の増加率 B_1 を確認する。ファクター 2^m を導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギー $C_{(m,s)}$ の観察では検出しきらなかった写像の過度の歪みを検出するために用いることができる。

【0064】なお実験では、副写像 $f_{(m,s)}$ を計算

【数 25】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (\text{式 2 5})$$

と仮定すれば、

【数 26】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (\text{式 2 6})$$

が成り立つ。このとき $k \neq -3$ であれば、

【数 27】

$$(\text{式 2 7})$$

する際、もし λ が $0, 1$ を越えたら $f_{(m,s)}$ の計算は止めて $f_{(m,s+1)}$ の計算に移行した。 $\lambda > 0, 1$

のとき、画素の輝度 255 レベル中のわずか「3」の違いが副写像の計算に影響したためであり、 $\lambda > 0, 1$ のとき正しい結果を得ることは困難だったためである。

【0065】 [1. 4. 2] ヒストグラム $h(l)$

$C_{(m,s)}$ の検査はヒストグラム $h(l)$ に依存しない。全単射及びその第3の条件の検査の際、 $h(l)$ に影響を受けうる。実際に $(\lambda, C_{(m,s)})$ をプロットすると、 k は通常1付近にある。実験では $k=1$ を用い、 $B_0 \lambda^2$ と $B_1 \lambda^2$ を検査した。仮に k の本当の値が1未満であれば、 $B_0 \lambda^2$ と $B_1 \lambda^2$ は定数にならず、ファクター $\lambda^{(1-k)/2}$ に従って徐々に増加する。 $h(l)$ が定数であれば、例えばファクターは $\lambda^{1/2}$ である。しかし、こうした差は閾値 B

threshold を正しく設定することによって吸収することができる。

【0066】ここで次式のごとく始点画像を中心が (x_0, y_0) 、半径 r の円形のオブジェクトであると仮定する。

【数 32】

$$p(i,j) = \begin{cases} \frac{2\pi}{r} c \sqrt{(i-x_0)^2 + (j-y_0)^2} & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式 3 2})$$

一方、終点画像は、次式のごとく中心 (x_1, y_1) 、半径が r のオブジェクトであるとする。

【数 33】

$$q(i,j) = \begin{cases} \frac{2\pi}{r} c \sqrt{(i-x_1)^2 + (j-y_1)^2} & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式 3 3})$$

ここで $c(x)$ は $c(x) = x^k$ の形であるとする。中心 (x_0, y_0) 及び (x_1, y_1) が十分遠い場合、ヒストグラム $h(l)$ は次式の形となる。

【数 34】

$$h(l) \propto r l^k \quad (k \neq 0) \quad (\text{式 3 4})$$

$k=1$ のとき、画像は背景に埋め込まれた鮮明な境界線

を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。 $k = -1$ のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従つて暗くなる。一般的のオブジェクトはこれらふたつのタイプのオブジェクトの中間にあると考えてもさして一般性を失わない。したがって、 k は $-1 \leq k \leq 1$ として大抵の場合をカバーでき、式 27 が一般に減少関数であることが保障される。

【0067】なお、式 34 からわかるように、 r は画像の解像度に影響されること、すなわち r は 2 m に比例することに注意すべきである。このために [1. 4. 1]においてファクター 2 m を導入した。

【0068】[1. 4. 3] η の動的決定

パラメータ η も同様の方法で自動決定できる。はじめに $\eta = 0$ とし、最も細かい解像度における最終的な写像 $f^{(n)}$ 及びエネルギー $C_f^{(n)}$ を計算する。つづいて、 η をある値 $\Delta\eta$ だけ増加させ、再び最も細かい解像度における最終写像 $f^{(n)}$ 及びエネルギー $C_f^{(n)}$ を計算し直す。この過程を最適値が求まるまで続ける。 η は写像の剛性を示す。次式の重みだからである。

【0069】

【数 35】

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式 35)

$$V(q_{f^{(m,s)}(i,j)}^{(m,s)})$$

における輝度を持つ $f^{(m,s)}$ が提供される。つまりスーパーサンプリングが行われる。実験では、 f

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)})$$

は、

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2$$

(式 36)

によって与えられた。

【0072】[1. 6] 各画像の画素の輝度の正規化始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー $C_f^{(m,s)}$ が大きくなりすぎ、正しい評価がしづらいためである。

【0073】例えば、人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふたつの顔の間の副写像を計算するために、まず副画像

η が 0 のとき、 $D^{(n)}$ は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 η が非常に大きな値のとき、 $D^{(n)}$ は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 η の値が 0 から次第に増えるとき、後述のごとく $C_f^{(n)}$ は徐々に減少する。しかし η の値が最適値を越えると、図 4 に示すとおり、エネルギーは増加し始める。同図の X 軸は η 、Y 軸は $C_f^{(n)}$ である。

【0070】この方法で $C_f^{(n)}$ を最小にする最適な η の値を得ることができる。しかし、 λ の場合に比べていろいろな要素が計算に影響する結果、 $C_f^{(n)}$ は小さく揺らぎながら変化する。 λ の場合は、入力が微小量変化するたびに副写像を 1 回計算しなおすだけだが、 η の場合はすべての副写像が計算しなおされるためである。このため、得られた $C_f^{(n)}$ の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つかれば、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【0071】[1. 5] スーパーサンプリング

画素間の対応関係を決定する際、自由度を増やすために、 $f^{(m,s)}$ の値域を $R \times R$ に拡張することができる (R は実数の集合)。この場合、終点画像の画素の輝度が補間され、非整数点、

【数 36】

(式 36)

(m, s) は整数及び半整数値をとることが許され、

【数 37】

(式 37)

【数 38】

(式 38)

を正規化する。すなわち、最も暗い画素の輝度を 0、最も明るいそれを 255 に設定し、他の画素の輝度は線形補間によって求めておく。

【0074】[1. 7] インプリメンテーション

始点画像のスキャンに従って計算がリニアに進行する帰納的方法を用いる。始めに、1 番上の左端の画素 $(i, j) = (0, 0)$ について $f^{(m,s)}$ の値を決定する。次に i を 1 ずつ増やしながら各 $f^{(m,s)}(i, j)$ の値を決定する。 i の値が画像の幅に到達したとき、 j の値を 1 増やし、 i を 0 に戻す。以降、始点画像のスキャンに伴い $f^{(m,s)}(i,$

j) を決定していく。すべての点について画素の対応が決まれば、ひとつの写像 $f(m, s)$ が決まる。

【0075】ある $p(i, j)$ について対応点 $q_f(i, j)$ が決まれば、つぎに $p(i, j+1)$ の対応点 $q_f(i, j+1)$ が決められる。この際、 $q_f(i, j+1)$ の位置は全単射条件を満たすために、 $q_f(i, j)$ の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに $(0, 0)$ が最も優先される状態がつくと、求められる最終の写像に余計な偏向が加わる。本前提技術ではこの状態を回避するために、 $f(m, s)$ を以下の方法で決めていく。

【0076】まず $(s \bmod 4)$ が0の場合、 $(0, 0)$ を開始点とし i 及び j を徐々に増やしながら決めていく。 $(s \bmod 4)$ が1の場合、最上行の右端点を開始点とし、 i を減少、 j を増加させながら決めていく。 $(s \bmod 4)$ が2のとき、最下行の右端点を開始点とし、 i 及び j を減少させながら決めていく。 $(s \bmod 4)$ が3の場合、最下行の左端点を開始点とし、 i を増加、 j を減少させながら決めていく。解像度

$$W = \vec{A} \times \vec{B}$$

ただしここで、

$$\vec{A} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{f(m,s)}^{(m,s)}(i+1,j-1)}$$

【数41】

$$\vec{B} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{f(m,s)}^{(m,s)}(k,l)}$$

(式41)

である（ここでベクトルは三次元ベクトルとし、 z 軸は直交右手座標系において定義される）。もし W が負であれば、その候補については $D(m, s)_{(k, l)}$ に ϕ を掛けることによってペナルティを与え、できるかぎり選択しないようとする。

【0080】図5(a)、図5(b)はこの条件を検査する理由を示している。図5(a)はペナルティのない候補、図5(b)はペナルティがある候補をそれぞれ表す。隣接画素 $(i, j+1)$ に対する写像 $f(m, s)_{(i, j+1)}$ を決定する際、 W の z 成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、 $q_f(m, s)_{(k, l)}$ は隣接する四辺形の境界線を越えるためである。

【0081】[1. 7. 1] 副写像の順序

インプリメンテーションでは、解像度レベルが偶数のときは $\sigma(0) = 0, \sigma(1) = 1, \sigma(2) = 2, \sigma$

が最も細かい第 n レベルには副写像という概念、すなはちパラメータ s が存在しないため、仮に $s = 0$ 及び $s = 2$ であるとしてふたつの方向を連続的に計算した。

【0077】実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補 (k, l) の中からできる限り全単射条件を満たす $f(m, s)_{(i, j)}$ ($m = 0, \dots, n$) の値を選んだ。第3の条件を破る候補のエネルギー $D(k, l)$ には ϕ を掛け、一方、第1または第2の条件を破る候補には ψ を掛ける。今回は $\phi = 2, \psi = 100000$ を用いた。

【0078】前述の全単射条件のチェックのために、実際の手続として $(k, l) = f(m, s)_{(i, j)}$ を決定する際に以下のテストを行った。すなはち $f(m, s)_{(i, j)}$ の相続四辺形に含まれる各格子点 (k, l) に対し、次式の外積の z 成分が0以上になるかどうかを確かめる。

【0079】

【数39】

(式39)

【数40】

(式40)
 $(3) = 3, \sigma(4) = 0$ を用い、奇数のときは $\sigma(0) = 3, \sigma(1) = 2, \sigma(2) = 1, \sigma(3) = 0, \sigma(4) = 3$ を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は4種類であり、 s は0~3のいずれかである。しかし、実際には $s = 4$ に相当する処理を行った。その理由は後述する。

【0082】[1. 8] 補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面における正方形 p

$(i, j) p(i+1, j) p(i, j+1) p(i+1, j+1)$ が終点画像平面上の四辺形 $q_f(i, j) q_f(i+1, j) q_f(i, j+1) q_f(i+1, j+1)$ に射影されると仮定する。簡単のため、画像間の距離を1とする。始点画像平面からの距離が t ($0 \leq t \leq 1$) である中間画像の画素 $r(x, y, t)$ ($0 \leq x \leq N-1, 0 \leq y \leq M-1$) は以下の要領で求められる。まず画素 $r(x, y, t)$ の位置 (ただし $x, y, t \in R$) を次式で求める。

【0083】

【数42】

$$\begin{aligned}
 (x, y) &= (1 - dx)(1 - dy)(1 - t)(i, j) + (1 - dx)(1 - dy)t f(i, j) \\
 &+ dx(1 - dy)(1 - t)(i + 1, j) + dx(1 - dy)t f(i + 1, j) \\
 &+ (1 - dx)dy(1 - t)(i, j + 1) + (1 - dx)dyt f(i, j + 1) \\
 &+ dxdy(1 - t)(i + 1, j + 1) + dxdytf(i + 1, j + 1)
 \end{aligned}$$

(式 4 2)

つづいて $r(x, y, t)$ における画素の輝度が次の式を用いて決定される。

【0084】

【数 4 3】

$$\begin{aligned}
 V(r(x, y, t)) &= (1 - dx)(1 - dy)(1 - t)V(p_{(i,j)}) + (1 - dx)(1 - dy)tV(q_{f(i,j)}) \\
 &+ dx(1 - dy)(1 - t)V(p_{(i+1,j)}) + dx(1 - dy)tV(q_{f(i+1,j)}) \\
 &+ (1 - dx)dy(1 - t)V(p_{(i,j+1)}) + (1 - dx)dytV(q_{f(i,j+1)}) \\
 &+ dxdy(1 - t)V(p_{(i+1,j+1)}) + dxdytfV(q_{f(i+1,j+1)})
 \end{aligned}$$

(式 4 3)

ここで $d x$ 及び $d y$ はパラメータであり、0から1まで変化する。

【0085】 [1. 9] 拘束条件を課したときの写像
いままでは拘束条件がいっさい存在しない場合の写像の決定を述べた。しかし、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束条件としたうえで写像を決定することができる。

【0086】 基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、かかる後、写像 f を正確に計算する。

$$p_{(i_0, j_0)}, p_{(i_1, j_1)}, \dots,$$

$$p_{(i_{n_s-1}, j_{n_s-1})}$$

(式 4 4)

【数 4 5】

を特定するとき、以下の値を決める。

$$\begin{aligned}
 F^{(n)}(i_0, j_0) &= (k_0, l_0), \\
 F^{(n)}(i_1, j_0) &= (k_1, l_1), \dots,
 \end{aligned}$$

$$F^{(n)}(i_{n_s-1}, j_{n_s-1}) = (k_{n_s-1}, l_{n_s-1})$$

(式 4 5)

始点画像の他の画素の変位量は、 $p_{(i_h, j_h)}$ ($h = 0, \dots, n_s - 1$) の変位に重み付けをして求められる平均である。すなわち画素 $p_{(i_h, j_h)}$ は、終点画像

の以下の画素に射影される。

【0089】

【数 4 6】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{h=n_s-1} (k_h - i_h, l_h - j_h) weight_h(i, j)}{2^{n_m}}$$

(式 4 6)

ただしここで、

【数 4 7】

$$weight_h(i, j) = \frac{1 / ||(i_h - i, j_h - j)||^2}{total weight(i, j)} \quad (式 4 7)$$

【数 4 8】

$$total weight(i, j) = \sum_{h=0}^{h=n_s-1} 1 / ||(i_h - i, j_h - j)||^2 \quad (式 4 8)$$

とする。

【0090】つづいて、 $F^{(m)}$ に近い候補写像 f がより少ないエネルギーを持つように、その写像 f のエネルギー

$$D_{(i,j)}^{(m,s)} = E_{0(i,j)}^{(m,s)} + \eta E_{1(i,j)}^{(m,s)} + \kappa E_{2(i,j)}^{(m,s)}$$

(式49)

である。ただし、

$$E_{2(i,j)}^{(m,s)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i,j) - f^{(m,s)}(i,j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i,j) - f^{(m,s)}(i,j)\|^2, & \text{otherwise} \end{cases}$$

(式50)

【数50】

であり、 $\kappa, \rho \geq 0$ とする。最後に、前述の写像の自動計算プロセスにより、 f を完全に決定する。

【0091】ここで、 $f^{(m,s)}(i,j)$ が $F^{(m)}(i,j)$ に十分近いとき、つまりそれらの距離が、

【数51】

$$\left[\frac{\rho^2}{2^{2(n-m)}} \right] \quad (\text{式51})$$

以内であるとき、 $E_{2(i,j)}^{(m,s)}$ が 0 になることに注意すべきである。そのように定義した理由は、各 $f^{(m,s)}(i,j)$ が $F^{(m)}(i,j)$ に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

【0092】[2] 具体的な処理手順

[1] の各要素技術による処理の流れを説明する。図6は前提技術の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い (S1)、つづいて始点画像と終点画像のマッチングをとる (S2)。ただし、S2は必須ではなく、S1で得られた画像の特徴をもとに画像認識などの処理を行ってよい。

【0093】図7は図6のS1の詳細を示すフローチャートである。ここではS2で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い (S10)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い (S11)、一連の終点階層画像を得る。ただし、S10とS11の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

【0094】図8は図7のS10の詳細を示すフローチャートである。との始点画像のサイズは $2^n \times 2^n$ とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ m を n にセットする (S100)。つづいて第 m レベルの画像 $p^{(m,0)}, p^{(m,1)}, p^{(m,2)}, p^{(m,3)}$ から特異点フィルタを用いて特異点を検出し (S101)、それぞれ第 $m-1$ レベルの画像 $p^{(m-1,0)}, p^{(m-1,1)}, p^{(m-1,2)}, p^{(m-1,3)}$ を生成する (S102)。ここでは $m=n$ であるため、 $p^{(m,0)}=p^{(m,1)}=p^{(m,2)}=p^{(m,3)}=p^{(n)}$ であり、ひとつの始点画像から4種類の副画像が生成される。

ギー D $(m, s)(i, j)$ を変更する。正確には、D $(m, s)(i, j)$ は、

【数49】

ベルの画像 $p^{(m,0)}, p^{(m,1)}, p^{(m,2)}, p^{(m,3)}$ から特異点フィルタを用いて特異点を検出し (S101)、それぞれ第 $m-1$ レベルの画像 $p^{(m-1,0)}, p^{(m-1,1)}, p^{(m-1,2)}, p^{(m-1,3)}$ を生成する (S102)。ここでは $m=n$ であるため、 $p^{(m,0)}=p^{(m,1)}=p^{(m,2)}=p^{(m,3)}=p^{(n)}$ であり、ひとつの始点画像から4種類の副画像が生成される。

【0095】図9は第 m レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の数値は各画素の輝度を示す。同図の $p^{(m,s)}$ は $p^{(m,0)} \sim p^{(m,3)}$ の4つの画像を象徴するもので、 $p^{(m-1,0)}$ を生成する場合には、 $p^{(m,s)}$ は $p^{(m,0)}$ であると考える。[1, 2] で示した規則により、 $p^{(m-1,0)}$ は例えば同図で輝度を記入したブロックについて、そこに含まれる4画素のうち「3」、 $p^{(m-1,1)}$ は「8」、 $p^{(m-1,2)}$ は「6」、 $p^{(m-1,3)}$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2^{m-1} \times 2^{m-1}$ になる。

【0096】つづいて m をデクリメントし (図8の S103)、 m が負になっていないことを確認し (S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0レベルの副画像が生成された時点で S10 が終了する。第0レベルの副画像のサイズは 1×1 である。

【0097】図10は S10 によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが4つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7の S11 にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6の S1 による処理が完了する。

【0098】前提技術では、図6の S2 に進むためにマッチング評価の準備をする。図11はその手順を示している。同図のごとく、まず複数の評価式が設定される

(S30)。[1. 3. 2. 1]で導入した画素に関するエネルギーC_f(m, s)と[1. 3. 2. 2]で導入した写像の滑らかさに関するエネルギーD_f(m, s)_fがそれである。つぎに、これらの評価式を統合して総

$$\sum \sum (\lambda C_f(m, s)_{(i, j)} + \eta E_0(m, s)_{(i, j)}) \quad (式52)$$

となる。ただし、総和はi、jについてそれぞれ0、1…、2^{m-1}で計算する。以上でマッチング評価の準備が整う。

【0099】図12は図6のS2の詳細を示すフローチャートである。[1]で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしでとられる。画像の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優れたものになる。

【0100】図12のごとく、まず係数パラメータηを0、レベルパラメータmを0に設定する(S20)。つづいて、始点階層画像中の第mレベルの4つの副画像と終点階層画像中の第mレベルの4つの副画像のそれぞれの間でマッチングを計算し、それぞれ全単射条件を満たし、かつエネルギーを最小にするような4種類の副写像f_f(m, s)(s=0, 1, 2, 3)を求める(S21)。全単射条件は[1. 3. 3]で述べた相続四辺形を用いて検査される。この際、式17、18が示すように、第mレベルにおける副写像は第m-1レベルのそれに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いまm=0であってそれより粗いレベルはないが、この例外的な処理は図13で後述する。

【0101】一方、同一レベル内における水平的参照も行われる。[1. 3. 3]の式20のごとく、f_f(m, 3)はf_f(m, 2)に、f_f(m, 2)はf_f(m, 1)に、f_f(m, 1)はf_f(m, 0)に、それぞれ類似するように決める。その理由は、特異点の種類が違っても、それらがもともと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式20からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが良好とみなされる。

【0102】なお、最初に決めるべきf_f(m, 0)については同一のレベルで参照できる副写像がないため、式19に示すごとくひとつ粗いレベルを参照する。ただし、実験ではf_f(m, 3)まで求まつた後、これを拘束条件としてf_f(m, 0)を一回更新するという手続をとった。これは式20にs=4を代入し、f_f(m, 4)を

合評価式を立てる(S31)。[1. 3. 2. 3]で導入した総エネルギーλC_f(m, s)_f+D_f(m, s)_fがそれであり、[1. 3. 2. 2]で導入したηを用いれば、

$$\sum \sum (\lambda C_f(m, s)_{(i, j)} + \eta E_0(m, s)_{(i, j)}) + E_1(m)$$

新たなf_f(m, 0)とすることに等しい。f_f(m, 0)とf_f(m, 3)の関連度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では[1. 7. 1]に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの関連度を密接に保つ趣旨である。また、処理の開始点に依存する偏向を回避するために、sの値にしたがって開始点の位置を変える点は[1. 7]で述べたとおりである。

【0103】図13は第0レベルにおいて副写像を決定する様子を示す図である。第0レベルでは各副画像がただひとつの画素で構成されるため、4つの副写像f_f(0, s)はすべて自動的に単位写像に決まる。図14は第1レベルにおいて副写像を決定する様子を示す図である。第1レベルでは副画像がそれぞれ4画素で構成される。同図ではこれら4画素が実線で示されている。いま、p(1, s)の点xの対応点をq(1, s)の中で探すとき、以下の手順を踏む。

- 【0104】1. 第1レベルの解像度で点xの左上点a、右上点b、左下点c、右下点dを求める。
2. 点a～dがひとつ粗いレベル、つまり第0レベルにおいて属する画素を探す。図14の場合、点a～dはそれぞれ画素A～Dに属する。ただし、画素A～Cは本来存在しない仮想的な画素である。
3. 第0レベルすでに求まっている画素A～Dの対応点A'～D'をq(1, s)の中にプロットする。画素A'～C'は仮想的な画素であり、それぞれ画素A～Cと同じ位置にあるものとする。
4. 画素Aの中の点aの対応点a'が画素A'の中にあるとみなし、点a'をプロットする。このとき、点aが画素Aの中で占める位置（この場合、右下）と、点a'が画素A'の中で占める位置が同じであると仮定する。
5. 4と同様の方法で対応点b'～d'をプロットし、点a'～d'で相続四辺形を作る。
6. 相続四辺形の中でエネルギーが最小になるよう、点xの対応点x'を探す。対応点x'の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図14の場合、4つの画素がすべて候補になる。

【0105】以上がある点xの対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第2レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図3に示すように画素A'～D'の間隔が空いていく状況が発生する。

【0106】こうして、ある第mレベルの4つの副写像が決まれば、mをインクリメントし(図12のS22)、mがnを超えていないことを確かめて(S23)、S21に戻る。以下、S21に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後にS21に戻ったときに第nレベルの写像 $f^{(n)}$ を決める。この写像は $\eta=0$ に関して定まったものであるから、 $f^{(n)}(\eta=0)$ と書く。

【0107】つぎに異なる η に関する写像も求めるべく、 η を $\Delta\eta$ だけシフトし、mをゼロクリアする(S24)。新たな η が所定の探索打切り値 η_{max} を超えていないことを確認し(S25)、S21に戻り、今回の η に関して写像 $f^{(n)}(\eta=\Delta\eta)$ を求める。この処理を繰り返し、S21で $f^{(n)}(\eta=i\Delta\eta)$ ($i=0, 1, \dots$)を求めていく。 η が η_{max} を超えたときS26に進み、後述の方法で最適な $\eta=\eta_{opt}$ を決定し、 $f^{(n)}(\eta=\eta_{opt})$ を最終的に写像 $f^{(n)}$ とする。

【0108】図15は図12のS21の詳細を示すフローチャートである。このフローチャートにより、ある定まった η について、第mレベルにおける副写像が決まる。副写像を決める際、前提技術では副写像ごとに最適な λ を独立して決める。

【0109】同図のごとく、まずsと λ をゼロクリアする(S210)。つぎに、そのときの λ について(および暗に η について)エネルギーを最小にする副写像 $f^{(m, s)}$ を求め(S211)、これを $f^{(m, s)}(\lambda=0)$ と書く。異なる λ に関する写像も求めるべく、 λ を $\Delta\lambda$ だけシフトし、新たな λ が所定の探索打切り値 λ_{max} を超えていないことを確認し(S213)、S211に戻り、以降の繰り返し処理で $f^{(m, s)}(\lambda=i\Delta\lambda)$ ($i=0, 1, \dots$)を求める。 λ が λ_{max} を超えたときS214に進み、最適な $\lambda=\lambda_{opt}$ を決定し、 $f^{(m, s)}(\lambda=\lambda_{opt})$ を最終的に写像 $f^{(m, s)}$ とする(S214)。

【0110】つぎに、同一レベルにおける他の副写像を求めるべく、 λ をゼロクリアし、sをインクリメントする(S215)。sが4を超えていないことを確認し(S216)、S211に戻る。s=4になれば上述のごとく $f^{(m, 3)}$ を利用して $f^{(m, 0)}$ を更新し、そのレベルにおける副写像の決定を終了する。

【0111】図16は、あるmとsについて λ を変えながら求められた $f^{(m, s)}(\lambda=i\Delta\lambda)$ ($i=0, 1, \dots$)に対応するエネルギー $C^{(m, s)}$ の挙動を示す図である。[1. 4]で述べたとおり、 λ が増加すると通常 $C^{(m, s)}$ は減少する。しかし、 λ が最適値を超えると $C^{(m, s)}$ は増加に転じる。そこで本前提技術では $C^{(m, s)}$ が極小値をとるときの λ を λ_{opt} と決める。同図のように $\lambda > \lambda_{opt}$ の範囲で再度 $C^{(m, s)}$ が小さくなってしまっても、その時点

ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 λ_{opt} は副写像ごとに独立して決めていく、最後に $f^{(n)}$ についてもひとつ定まる。

【0112】一方、図17は、 η を変えながら求められた $f^{(n)}(\eta=i\Delta\eta)$ ($i=0, 1, \dots$)に対応するエネルギー $C^{(n)}$ の挙動を示す図である。ここでも η が増加すると通常 $C^{(n)}$ は減少するが、 η が最適値を超えると通常 $C^{(n)}$ は増加に転じる。そこで $C^{(n)}$ が極小値をとるときの η を η_{opt} と決める。図17は図4の横軸のゼロ付近を拡大した図と考えてよい。 η_{opt} が決まれば $f^{(n)}$ を最終決定することができる。

【0113】以上、本前提技術によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先駆的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

【0114】なお、本前提技術について次のような変形技術も考えられる。

(1) 前提技術では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の2枚の画像間のマッチングをとる場合全般に利用できる。

【0115】たとえば2枚の画像間で、画素の輝度の差に関するエネルギー E_0 と画素の位置的なずれに関するエネルギー E_1 のふたつを評価式とし、これらの線形和 $E_{tot} = \alpha E_0 + E_1$ を総合評価式とする。この総合評価式の極値付近に注目して α を自動決定する。つまり、いろいろな α について E_{tot} が最小になるような写像を求める。それらの写像のうち、 α に関して E_1 が極小値をとるときの α を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

【0116】これ以外にも評価式の設定にはいろいろな方法があり、例えば $1/E_1$ と $1/E_2$ のように、評価結果が良好なほど大きな値をとるものを探用してもよい。総合評価式も必ずしも線形和である必要はなく、n乗和($n=2, 1/2, -1, -2$ など)、多項式、任意の関数などを適宜選択すればよい。

【0117】パラメータも、 α のみ、前提技術のごとく η と λ のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが3以上の場合はひとつずつ変化させて決めていく。

【0118】(2) 本前提技術では、総合評価式の値が

最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である $C(m, s)_f$ が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によっては単に総合評価式の最小値が最小になるようにパラメータを決めて効果的である。その場合、例えば $\alpha E_0 + \beta E_1$ を総合評価式とし、 $\alpha + \beta = 1$ なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてもよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

【0119】(3) 前提技術では各解像度レベルで4種類の特異点に関する4種類の副画像を生成した。しかし、当然4種類のうち1、2、3種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する $f(m, s)$ だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 s に関する計算量が減る効果がある。

【0120】(4) 本前提技術では特異点フィルタによってレベルがひとつ進むと画素が $1/4$ になった。例えば 3×3 で1ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は $1/9$ になる。

【0121】(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像を変換する。それ以外の方法として、RGBの各成分について副写像を計算してもよい。

【0122】[映像効果に関する実施の形態] 以上の前提技術を利用した映像効果装置を説明する。この装置は、動画をデジタルカメラに効率的に記録するための基礎技術を利用しておらず、その説明のあと、図24以降でデジタルカメラへの適用を述べる。図18は、キーフレームである第1画像I1、第2画像I2を示し、それらの間で、ある画素 $p_1(x_1, y_1)$ と $p_2(x_2, y_2)$ が対応している。これらの対応は前提技術で求められている。

【0123】図19は、第1画像I1上にメッシュを切り、そのメッシュを構成するポリゴンの第2画像I2上の対応位置を示す。いま第1画像I1上で注目するポリゴンR1は4個の格子点A、B、C、Dによって決まっている。このポリゴンR1を「起点ポリゴン」と名付ける。これら4個の格子点A、B、C、Dは、図18に示したごとく、それぞれ第2画像I2上に対応する点A'、B'、C'、D'を有し、これらの点で形成されるポリゴンR2を「終点ポリゴン」と名付ける。起点ポリゴンは一般に長方形であるが、終点ポリゴンは一般に四辺形である。いずれにせよ、この実施の形態では、画素単位で第1画像I1と第2画像I2の対応関係を記述せず、起点ポリゴンの格子点について、その対応画素を

記述する。その記述は対応点ファイルに書き込まれる。格子点に注目することにより、対応点ファイルの容量を非常に小さくすることができる。

【0124】対応点ファイルは第1画像I1と第2画像I2の中間画像を生成するために利用される。この点は前提技術で述べたごとく、対応点どうしの位置を補間することにより、任意の時間的位置の中間画像を生成できる。したがって、第1画像I1、第2画像I2、および対応点ファイルを保存しておくことで、それに2つの画像間のモーフィングや滑らかな動画を生成することができる。このため、動画の圧縮効果が得られる。

【0125】図20は対応点ファイルから格子点以外の点に関する対応関係を計算する方法を示す。対応点ファイルには格子点に関する情報しかないので、ポリゴン内部の点について対応情報を別途計算する必要がある。図20では、図19の起点ポリゴンR1の下半分に当たる三角形ABCと、同様に終点ポリゴンR2の下半分に当たる三角形A'B'C'との対応を示す。いま、起点ポリゴン側の三角形ABC内部の点Qが、線分ACを $t : (1-t)$ に内分し、その内分点と点Bの間を $s : (1-s)$ に内分する位置にあったとする。この点Qの終点ポリゴン側の三角形A'B'C'における対応点Q'は、線分A'C'を $t : (1-t)$ に内分し、その内分点と点B'の間を $s : (1-s)$ に内分する位置にあると考えればよい。すなわち、起点ポリゴンを三角形に分割し、その三角形に関するベクトルの内分の形式にて終点ポリゴン内部の点を求めればよい。ベクトル斜体で示せば、

$$BQ = (1-s) \{ (1-t) BA + t BC \}$$

であり、したがって、

$$B'Q' = (1-s) \{ (1-t) B'A' + t B'C' \}$$

当然ながら、同様の処理を起点ポリゴンR1の上半分に当たる三角形ACDと、同様に終点ポリゴンR2の上半分に当たる三角形A'C'D'との間についてもなすとする。

【0126】図21は、以上の処理手順を示す。まず図19のごとく、第1画像I1上に取られた格子点に関するマッチング結果を取得する(S10)。このとき、前提技術の画素単位のマッチングを実行し、その結果の中から格子点に当たる部分を抽出すればよい。ただし、前提技術を用いて他のマッチング技術、例えばオプティカルフローやブロックマッチングをもとに格子点に関するマッチング結果を特定してもよい。

【0127】つづいて、図19の右側のごとく、第2画像I2上に終点ポリゴンを定義する(S12)。以上で対応点ファイルが生成できるので、このファイルに第1画像I1、第2画像I2を識別するための情報を盛り込んで出力する(S14)。ふたつの画像および対応点ファイルは、任意の記録装置または媒体に保持され、また

は直接ネットワークや放送波に乗せて伝送されればよい。

【0128】図22は、対応点ファイルを用いて中間画像を生成する手順を示す。まず、第1画像I1、第2画像I2が読み込まれ(S20)、対応点ファイルが読み込まれる(S22)。つづいて、起点ポリゴン内の点と終点ポリゴン内の点の対応関係が図20の方法によって算出される(S24)。この時点で、画像内のすべての画素に関する対応関係を取得することができる。したがって、前提技術で述べたごとく、互いに対応し合う点の座標と色をu:(1-u)で内分すれば、第1画像I1と第2画像I2の間を時間的に(1-u):uで内分する位置における中間画像を生成することができる(S26)。なお、内挿補間だけでなく、外挿補間をしてもよい。

【0129】図23は以上の処理を行う映像効果装置10の構成を示す。この装置10は、外部の記憶装置や撮影カメラ等から第1画像I1、第2画像I2を取得する画像入力部12と、それらの画像に前提技術その他の技術によってマッチング計算を施すマッチングプロセッサ14と、それによって生成された対応点ファイルFを格納する対応点ファイル保持部16と、第1画像I1および第2画像I2から中間画像を生成する中間画像生成部18と、第1画像I1、第2画像I2および中間画像をタイミング調整しながらもとの動画に近い映像として表示する表示部20を含む。また、通信部22は、外部からの要望にしたがい、第1画像I1、第2画像I2および対応点ファイルFをネットワークその他の伝送インターフェースへ送出する。なお、マッチングプロセッサ14には、メッシュの大きさや格子点の位置などを示すメッシュ情報が入力されている。

【0130】以上の構成により、画像入力部12に入力された第1画像I1と第2画像I2はマッチングプロセッサ14へ送られる。マッチングプロセッサ14は、それらの画像間で画素単位のマッチング計算を行う。マッチングプロセッサ14はメッシュ情報をもとに対応点ファイルFを生成し、対応点ファイル保持部16へ出力する。

【0131】中間画像生成部18は、ユーザの求めに応じ、またはその他の要因で対応点ファイルFを読み出し、中間画像を生成する。この中間画像は表示部20へ送られ、そこで画像出力の時間調整が行われ、動画またはモーフィング画像が表示される。この動作からもわかるとおり、中間画像生成部18と表示部20は、本装置10とは別構成のリモート端末側にあってもよく、その場合、その端末が第1画像I1、第2画像I2、対応点ファイルFという比較的軽いデータを受け、自動的に動画を再生できる。

【0132】通信部22はそのための構成であり、ここではすでにリモート端末がネットワークの先で待ってい

ることを想定している。すなわち、通信部22は第1画像I1、第2画像I2、対応点ファイルFをネットワークや放送波に乗せて送り出し、リモート端末側で動画の表示が行われる。もちろんリモート端末は、表示ではなく、ストレージを目的としてもよい。なお本装置は、ネットワーク等の外部から通信部22を介して第1画像I1、第2画像I2およびそれらの対応点ファイルを入力し、これらを中間画像生成部18に渡して補間計算をし、中間画像を生成する用途もある。図23にはそのためのデータの経路Pが記述されている。

【0133】実施の形態の処理内容について実験を行ったところ、第1、第2画像として、例えば256×256程度の大きさを採用し、格子点を縦横10～数十画素おき程度に設定することにより、きわめて良好なモーフィングまたは動画圧縮効果が得られた。対応点ファイルの大きさは数キロから10キロバイト程度であり、画質の高さとデータ量の小ささが実現できることが確認できた。

【0134】[デジタルカメラの実施の形態] 図24は、図23の映像効果装置10をデジタルカメラ50に応用した構成を示す。以下、図23との相違を中心説明する。図24では、画像入力部12に変えて撮像部52が設けられ、これを制御するカメラコントローラ54が併設されている。また、対応点ファイル保持部16としてICカード制御部56が設けられ、これがICカード58へのデータの入出力を制御する。なお、第1画像I1、第2画像I2、対応点ファイルはすべてICカード制御部56を介してICカードへ書き可能に構成されている。通信部22は第1画像I1、第2画像I2、対応点ファイルをネットワーク、外部記憶装置、その他の外部伝送メディア等へ出力可能であり、同図ではICカード制御部56からデータを受け取る構成だが、もちろんそうでなく、通常のデータバスから受け取る構成でもよい。モード設定部70は、カメラコントローラ54による撮影モードを設定し、通常の静止画モード、動画モードのほか、実施の形態に特徴的な「簡易動画モード」の指定が可能である。

【0135】図25は撮像部52の一例を示す。映像はCCD(電荷結合素子)60で取得され、A/D変換器62でデジタル化され、記録前に前処理部64でホワイトバランス等の処理を受ける。最終形式の第1画像I1と第2画像I2は、ICカード58へ記録され、または直接マッチングプロセッサ14で処理を受ける。

【0136】図26は撮像部52の別例を示す。図25と異なり、一定の距離をおいてCCD60がふたつ設けられ、ステレオ画像の撮影が可能である。A/D変換器62、前処理部64はふたつのCCD60からの映像を時分割で処理する。ただし、これらも2系列設けてよい。

【0137】図24に戻り、カメラコントローラ14の

処理はいくとおりか考えられ、以下例を挙げる。

1. 単眼カメラとして動画を圧縮する用途

図25の単眼構成を採用するデジタルカメラ50において、1回の撮影動作において、撮像部52により、所定の時間間隔（以下単に「撮影間隔」という）で第1画像I1と第2画像I2を撮影する。このとき、デジタルカメラ50は、前述の簡易動画モード、すなわち静止画と動画の中間的な撮影モードにある。

【0138】このモードでユーザが撮影のためにリリースボタンを押すと、たとえば撮影間隔が1秒で2枚の画像が撮影される。この1秒の間に被写体が動いたり、ユーザが動いたりすることにより、第1画像I1と第2画像I2の間には差違が出る。この差違を埋めるべく、マッチングプロセッサ14が対応点ファイルを生成する。この対応点ファイルをもとに中間画像生成部18が中間画像を生成することで、1秒間にに対応する動画が生成できる。中間画像を多数生成することで再生動画の継続時間を1秒超に設定すれば、スローモーション画像も得られる。

【0139】生成された動画は液晶デバイス等による表示部20に表示され、ユーザは簡易動画の内容を確認できる。もちろん、表示部20は単に第1画像I1と第2画像I2のみを表示してもよい。いずれの場合も、対応点ファイルはICカード58に記録されるため、動画はデジタルカメラ50の外部の機器で表示することができる。その外部機器には中間画像生成部18同様の構成があるものとする。

【0140】このモードにおける撮影間隔をのばしていけば、当然ながら、より長い時間の動画を生成することができる。どの程度の時間を許すかは、画質との関係で決めればよく、ユーザが設定してもよい。なお、撮影時間もモード設定部70で決めてよい。

2. 単眼カメラとしてモーフィング画像を生成する用途

前記の撮影間隔がある程度以上になると、動画というよりも、次第にモーフィング画像の生成に近づいていく。したがって、それをデジタルカメラ50の仕様としてもよい。この場合さらに、撮影間隔という考え方をやめて、カメラコントローラ54の機能により、ユーザが第1画像I1と第2画像I2を選択できてもよい。それらの画像はすでに撮影した画像から選択してもよいし、ICカード58から入力された画像であってもよい。いずれにせよ、その場合、例えば全く関係のない画像間でもモーフィングができる。実験でも非常に良好なモーフィング画像が生成されている。

3. ステレオカメラとして多視点画像を生成する用途

図26の複眼構成を採用するデジタルカメラ50において、同時に2枚の画像を撮影し、マッチングプロセッサ14によって対応点ファイルを生成する。対応点ファイル

には、一方のCCD60で撮影された第1画像I1の各点について、他方のCCD60で撮影された第2画像I2における対応点を記述している。したがって、それらの画像の対応し合う点（以下「対応点対」という）の水平方向のずれから三角測量の原理でその点の奥行きが判明する。その結果、奥行きを強調するなどの処理により、特殊効果画像を生成することができる。

【0143】また、中間画像生成部18によれば、両CCD60間の視点画像が生成できる。外挿補間をすれば、CCD60間ではなく、デジタルカメラ50からある程度離れた視点からの画像を生成することができる。いろいろな視点を仮定することで、多視点画像も得られる。これがウォータースルー画像のもとにもなる。

【0144】なお、CCD60の一方または両方を着脱可能に設け、それらの空間距離をある程度とれるよう構成してもよい。その場合、ステレオカメラとしての性能が上がる。

【0145】以上、本発明に係るデジタルカメラの例を説明した。ここではパーソナルユースのカメラを中心に据えたが、本発明はプロフェッショナルユースのTVカメラや人工衛星に搭載されるカメラに適用されてもよい。

【0146】また、図23同様、デジタルカメラ50は、外部から通信部22やICカード58を介して第1画像I1、第2画像I2およびそれらの対応点ファイルを入力し、これらを中間画像生成部18に渡して補間計算をし、中間画像を生成してもよい。

【図面の簡単な説明】

【図1】 図1(a)とは図1(b)は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図1(c)と図1(d)は、ふたりの人物の顔に関して前提技術で求められる $p(5,0)$ の画像、図1(e)と図1(f)は、ふたりの人物の顔に関して前提技術で求められる $p(5,1)$ の画像、図1(g)と図1(h)は、ふたりの人物の顔に関して前提技術で求められる $p(5,2)$ の画像、図1(i)と図1(j)は、ふたりの人物の顔に関して前提技術で求められる $p(5,3)$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図2】 図2(R)はもとの四辺形を示す図、図2(A)、図2(B)、図2(C)、図2(D)、図2(E)はそれぞれ相続四辺形を示す図である。

【図3】 始点画像と終点画像の関係、および第mレベルと第m-1レベルの関係を相続四辺形を用いて示す図である。

【図4】 パラメータ η とエネルギー C_f の関係を示す図である。

【図5】 図5(a)、図5(b)は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図6】 前提技術の全体手順を示すフローチャートである。

【図7】 図6のS1の詳細を示すフローチャートである。

【図8】 図7のS10の詳細を示すフローチャートである。

【図9】 第mレベルの画像の一部と、第m-1レベルの画像の一部の対応関係を示す図である。

【図10】 前提技術で生成された始点階層画像を示す図である。

【図11】 図6のS2に進む前に、マッチング評価の準備の手順を示す図である。

【図12】 図6のS2の詳細を示すフローチャートである。

【図13】 第0レベルにおいて副写像を決定する様子を示す図である。

【図14】 第1レベルにおいて副写像を決定する様子を示す図である。

【図15】 図12のS21の詳細を示すフローチャートである。

【図16】 ある $f^{(m, s)}$ について λ を変えながら求められた $f^{(m, s)} (\lambda = i \Delta \lambda)$ に対応するエネルギー $C^{(m, s)} (\lambda)$ の挙動を示す図である。

【図17】 η を変えながら求められた $f^{(n)} (\eta = i \Delta \eta)$ ($i = 0, 1, \dots$) に対応するエネルギー $C^{(n)} (\eta)$ の挙動を示す図である。

【図18】 第1画像と第2画像間である画素の対応する様子を示す図である。

【図19】 第1画像上にとられた起点ポリゴンと第2

画像上にとられた終点ポリゴンの対応関係を示す図である。

【図20】 起点ポリゴン内の点に対応する終点ポリゴン内の点を求める手順を示す図である。

【図21】 実施の形態で対応点ファイルを生成する手順を示すフローチャートである。

【図22】 実施の形態で、対応点ファイルをもとに中間画像を生成する手順を示すフローチャートである。

【図23】 実施の形態に係る映像効果装置の構成図である。

【図24】 実施の形態に係るデジタルカメラの構成図である。

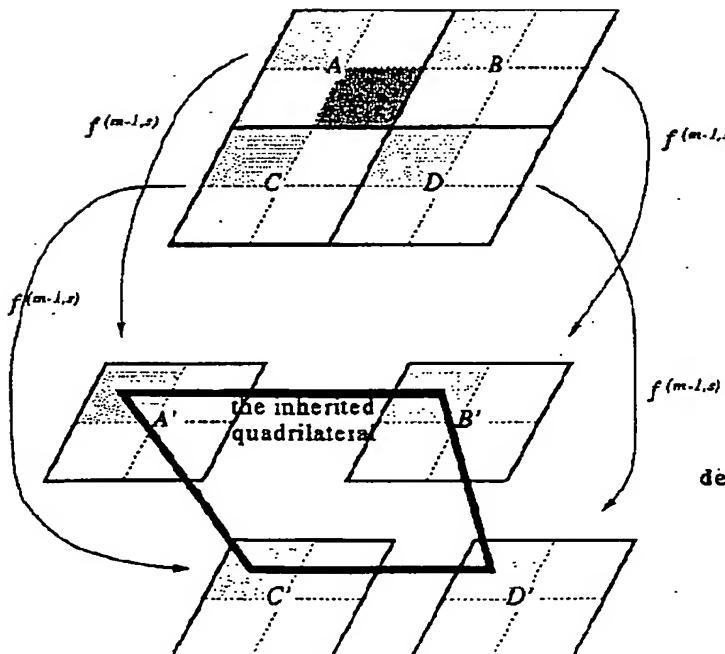
【図25】 図24に示すデジタルカメラの撮像部の構成図である。

【図26】 図24に示すデジタルカメラの撮像部の別の構成図である。

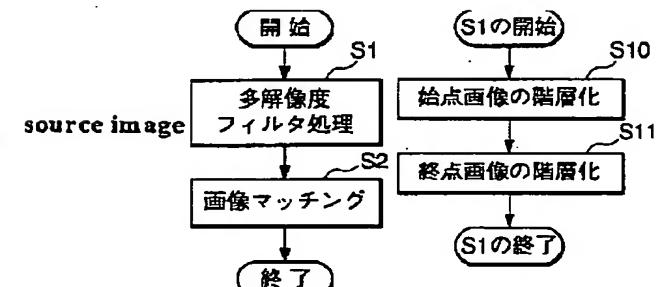
【符号の説明】

- 1 0 映像効果装置
- 1 2 画像データ入力部
- 1 4 マッチングプロセッサ
- 1 6 対応点ファイル保持部
- 1 8 中間画像生成部
- 2 0 表示部
- 2 2 通信部
- 5 0 デジタルカメラ
- 5 2 撮像部
- 5 4 カメラコントローラ
- 5 6 I C カード

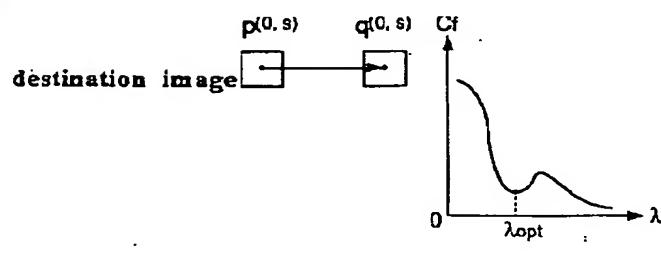
【図3】



【図6】

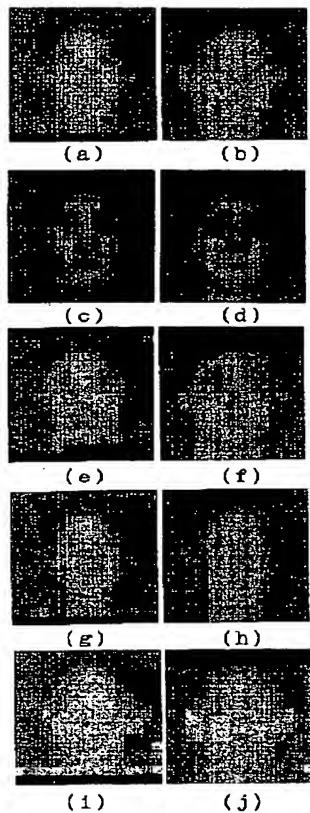


【図13】

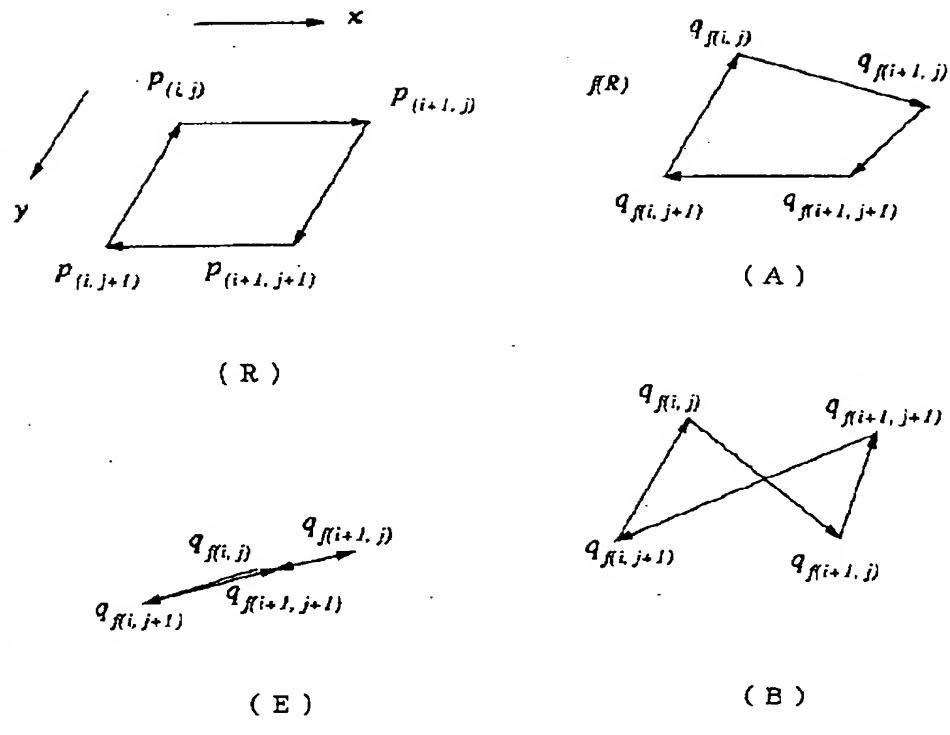


【図16】

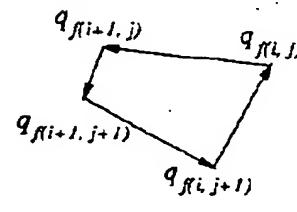
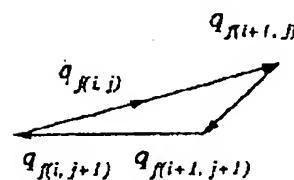
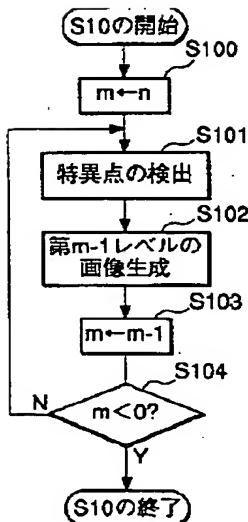
【図1】



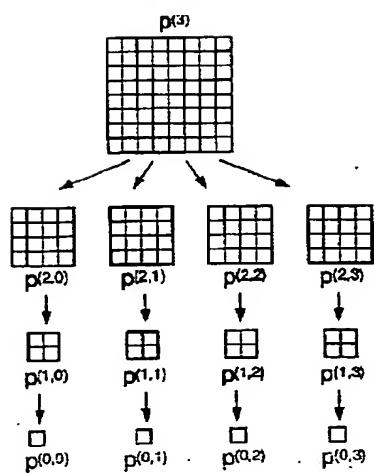
【図2】



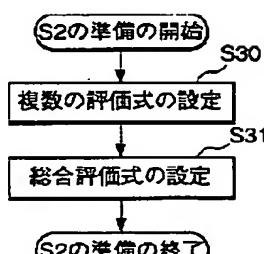
【図8】



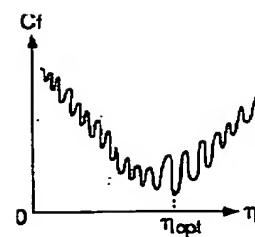
【図10】



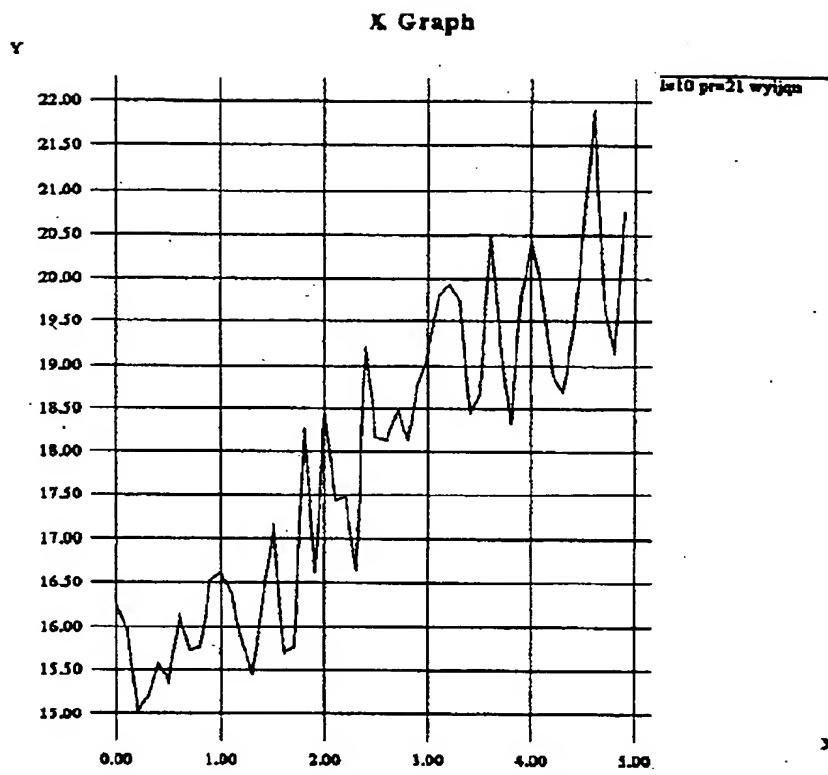
【図11】



【図17】

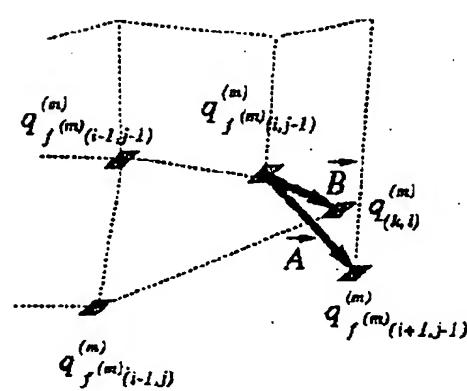
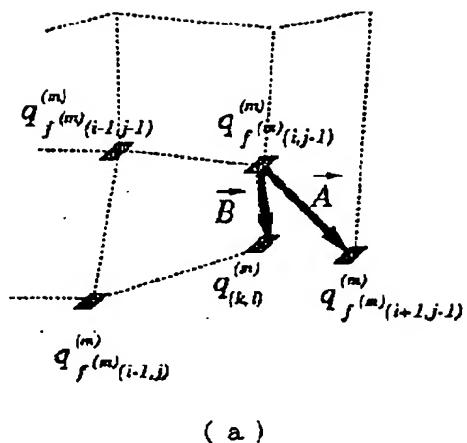


【図4】

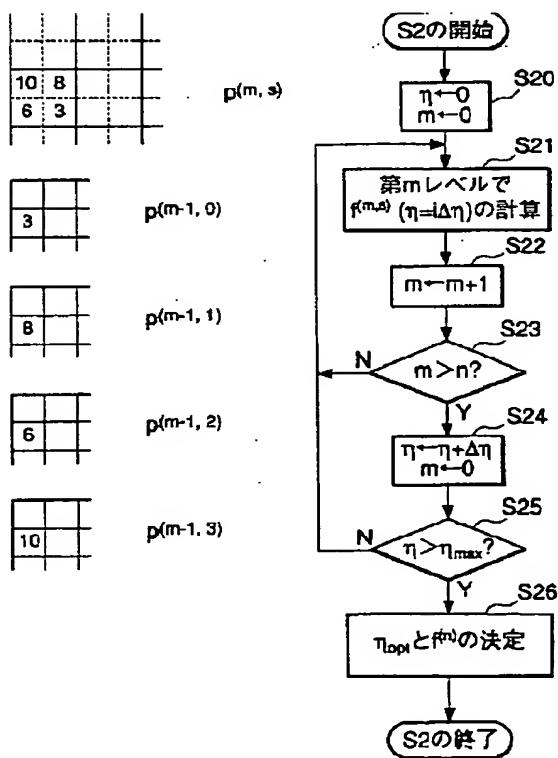


【図9】

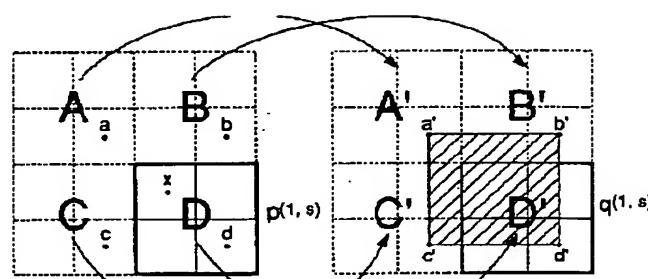
【図5】



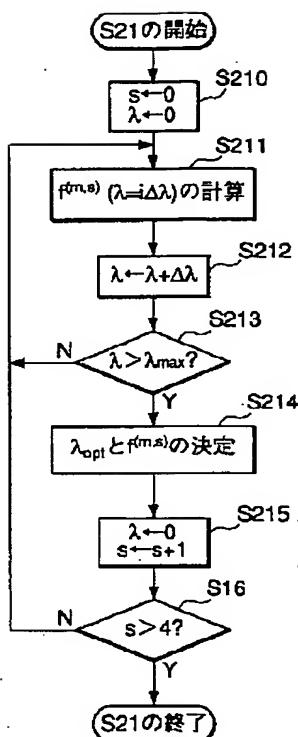
【図12】



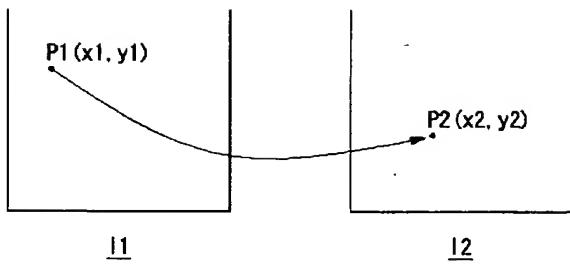
【図14】



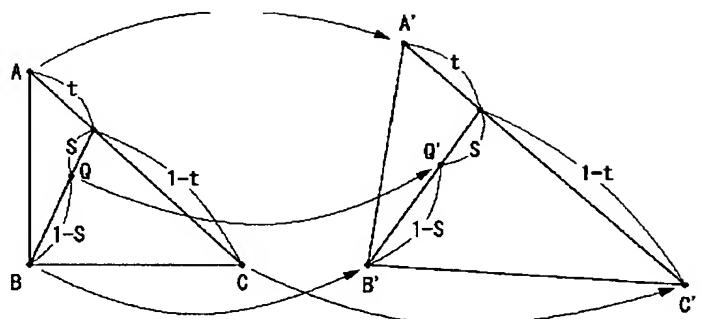
【図15】



【図18】



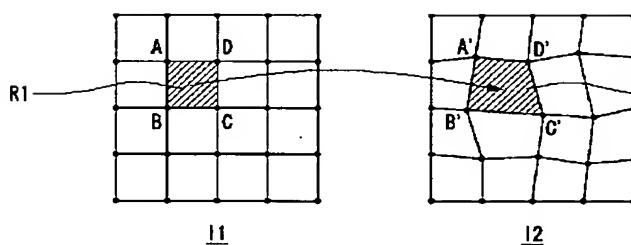
【図20】



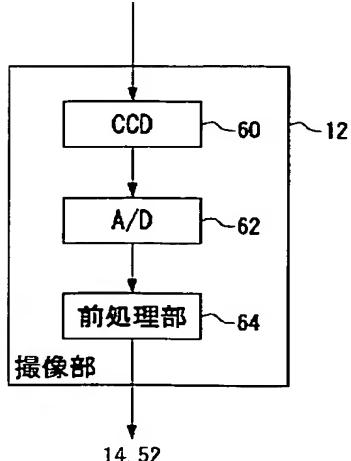
【図21】

【図22】

【図19】

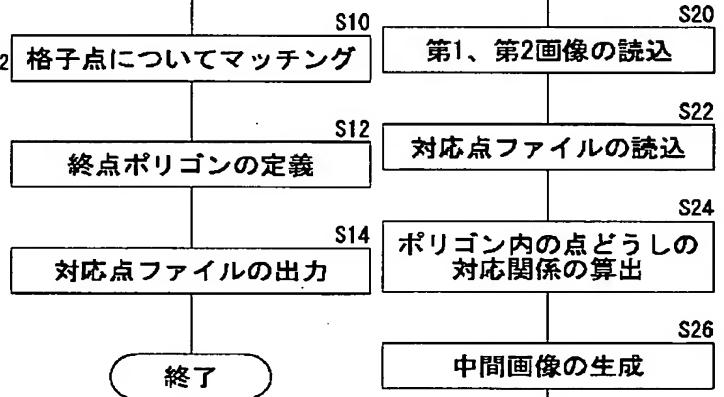


【図25】

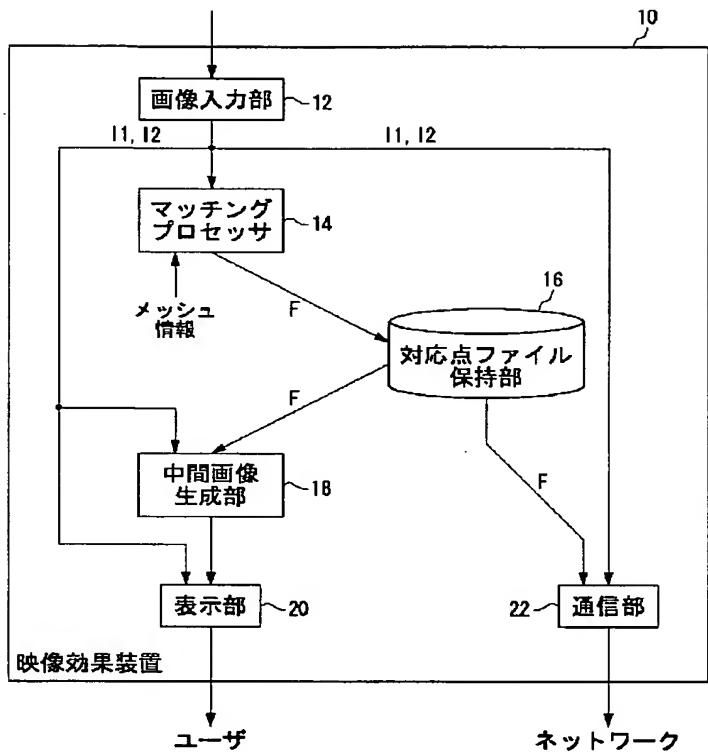


開始

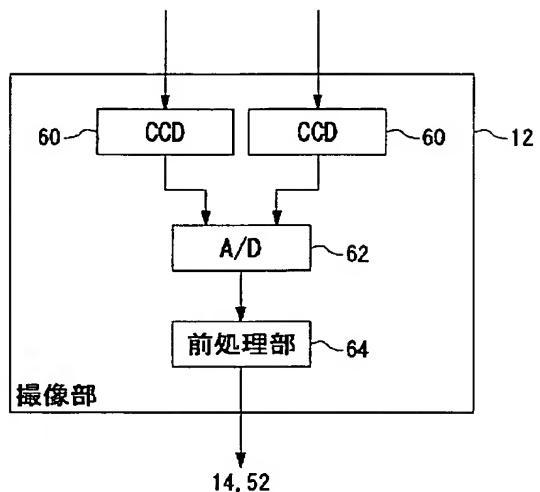
開始



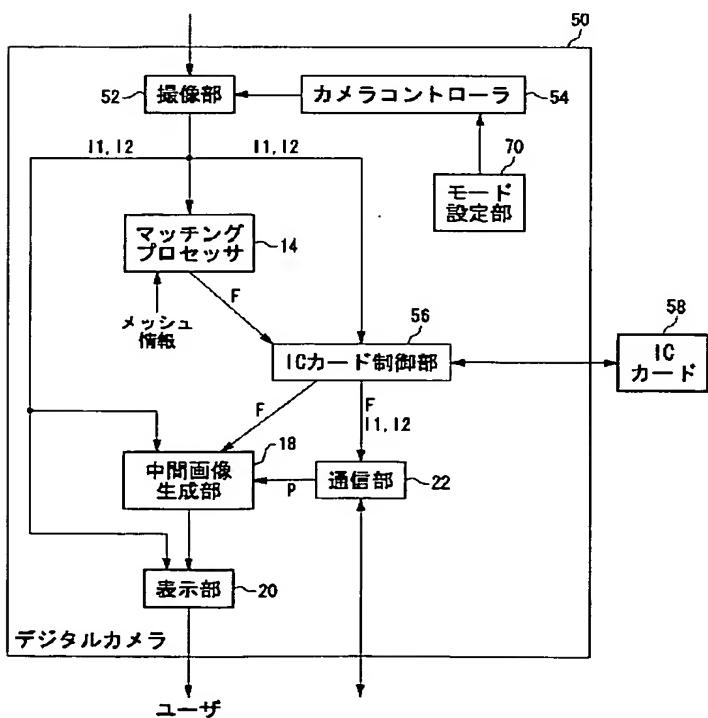
【図23】



【図26】



【図24】



フロントページの続き

(51) Int. C1. 7	識別記号	F I	マークコード(参考)
H O 4 N 13/02		H O 4 N 101:00	
// H O 4 N 101:00		5/91	J

F ターム(参考) 2H054 AA01
5C022 AA13 AC01 AC11 AC69 CA00
5C023 AA11 AA40 BA01 DA04 DA08
EA05
5C053 FA08 GB40 KA04 KA24 LA01
5C061 AA21 AB03 AB08 AB21